TECHNICAL MEMORANDUM

NO. 15

APPORTIONING HISTORICAL AGRICULTURAL STATISTICS USING GEO-GRIDDED DATA BASE ELEMENTS

UNITED STATES DEPARTMENT OF AGRICULTURE
FOREIGN AGRICULTURAL SERVICE
FOREIGN CROP CONDITION ASSESSMENT DIVISION
HOUSTON, TEXAS

Apportioning Historical Agricultural Statistics Using Geo-Gridded Data Base Elements	First Issue

APPROVED BY:

James R. Hickman

James R. Hickman, Director
Foreign Crop Condition Assessment Division

1. REASON FOR ISSUANCE

Document the results of this study for use by the Foreign Crop Condition Assessment Division (FCCAD) of the Foreign Agricultural Service (FAS). The results presented in this document serve as a basis for future research directed toward employing data elements stored in the geo-gridded data base to apportion historical agricultural statistics to the grid cell quadrant.

2. COVERAGE

This study evaluates allocation procedures which use current data elements of the geo-gridded data base.

3. ACKNOWLEDGMENT

Dr. Charles Perry of the ESS for his support in developing and analyzing the statistical results of the study. His contribution is hereby gratefully acknowledged.

4. CONCURRENCE

Bobby E. Spiers

Bobby E. Spiers, Chief
Commodity Analysis Branch
Foreign Crop Condition Assessment Division

5. PREPARED BY:

Ronald W. Willis

Ronald W. Willis
Foreign Crop Condition Assessment Division

9/14/82
DATE

SEP 23 1982

TABLE OF CONTENTS

TABLE OF CONTENTS

PART 1.0 INTRODUCTION

1.1 PURPOSE

The purpose of this study was to determine if data elements stored in the geo-gridded data base could be used to apportion historical agricultural area statistics to the 12.5 x 12.5 nautical mile grid cell quadrant.

1.2 SCOPE

The scope of this report was confined to the use of historical agricultural area statistics for one year for four crops (corn, wheat, soybeans, grain sorghum) from the province of Cordoba, Argentina and all grid cell quadrants which lie within the boundaries of that province. The scope was also limited to data available from the geo-gridded data base which are available for all countries.

1.3 BACKGROUND

The Foreign Crop Condition Assessment Division (FCCAD) of the Foreign Agricultural Service (FAS) is responsible for verifying and assessing the change in area planted to various crops in important foreign producing areas and reporting the results to FAS commodity analysts in Washington, D.C. The primary data source used by FCCAD analysts to assess the areal change is LANDSAT data. Because the FCCAD analysts monitor the majority of the crop region within a country and a great amount of time is required to analyze each LANDSAT image, certain high density or highly variable grid cell quadrants are selected to determine if a shift in cropping patterns has occurred. However, historical agricultural area statistics are available only at the state or country level for most countries. Therefore, to properly determine the crop mix in a particular grid cell quadrant, the state or country level historical statistics must be apportioned to the grid cell quadrant.

1.4 DATA SET

1.4.1 HISTORICAL AGRICULTURAL AREA STATISTICS

Province (state) and partido (county) level historical agricultural area statistics were available from Cordoba Province, Argentina for four crops for one year. The crops were corn, wheat, soybeans and grain sorghum with the year being 1978/79 for corn, soybeans and grain sorghum and 1977/78 for wheat. The partido level area statistics for each crop in each partido were initially apportioned equally to each quadrant in the partido. This procedure provided the initial figure for each crop in each quadrant for use in the correlation analysis.

1.4.2 AGRICULTURAL DENSITY

The agricultural density was determined by Dr. Fred Westin, South Dakota State University, using historical LANDSAT data and was divided into five categories based on percent of cultivated crop land. The categories and percent ranges were: 1=80-100 percent; 2=60-80 percent; 3=40-60 percent; 4=5-40 percent; and 5= less than 5 percent. Neither the categories nor the percent ranges provided a usable value with which to work. Consequently, the midpoint of each range was selected to replace the category. The categories were replaced as follows: 1 with 0.90; 2 with 0.70; 3 with 0.50; 4 with 0.225; and 5 with 0.025. Also stored in the data base with the categories were the reasons for the different rankings for each quadrant; e.g., "hve" was hilly with vegetation. Approximately 60 different reasons were available. The reasons were only used as an aide throughout the analysis.

1.4.3 SOILS DATA

The soil type chosen to represent the soils in the quadrant was the predominant soil in the quadrant. Associated with the soil types were five categories: texture, depth, drainage, salinity and slope. Each soil type received a rating from one to four in each of the five categories with one being the best and four the worst. The values in each category for each soil type were summed and then ranked with the lowest total receiving a ranking of one. Dummy values were also added for the five categories. The category, or categories, with the highest value was deemed to be the limiting factor(s) for that soil type. For example, a soil type might have these ratings: texture 1, depth 2, drainage 3, salinity 1, and slope 3. The limiting categories would be drainage and slope. Therefore, the dummy values would be: texture 0, depth 0, drainage 1, salinity 0, and slope 1.

1.4.4 AGRO-PHYSICAL UNITS

Agro-physical units (APU's) for Argentina were delineated by H. Edward Bulloch, FCCAD. The APU is a relatively homogeneous area similar in climate, soils, and crops. Each quadrant lies within only one APU. All APU's were grouped in similar regimes and then numbered within the regime.

1.5 APPROACH

To evaluate the apportioning procedure, accuracy was selected as the primary criterion for determining its general performance. Secondary criteria considered were efficiency and repeatability.

The accuracy criterion was selected as the primary criterion because if the apportioning procedure did not accurately determine the area to be apportioned to each grid cell quadrant, then the efficiency and repeatability criteria were moot questions to be considered. The apportioning procedures should accurately apportion the crop area to each grid cell quadrant as the analyst will use this as a basis for evaluating subsequent crop mixes in that grid cell quadrant.

PART 2.0 ANALYSIS

2.1 GRAPHS

Scatter plots were done to visually determine if any relationship existed between the crops and any of the other variables (Appendix 1). No clear relationship was indicated in any of the graphs. Therefore, no variables were deleted from consideration for use in the apportioning procedure.

2.2 CORRELATION AND APPORTIONMENT

2.2.1 County-level Statistics Equally Apportioned

The initial apportioning was to equally allocate the county-level statistics of each of the four crops to each grid cell quadrant in the province. This was accomplished by dividing the countylevel area statistics by the number of grid cell quadrants in the county and is in Appendix 2. A Spearman correlation was conducted and the results are presented in Table 1. No variable showed a high correlation with all four crops.

2.2.2 County-level Statistics Apportioned Using Ag Density

The next apportioning procedure utilized the quadrant level ag density to apportion the county-level statistics. The equation used is listed in Appendix 2. A Spearman correlation was conducted and the results are presented in Table 2. Ag density had the highest correlation with all four of the crops, but this should have occurred because it was used in the apportioning procedure. No other variables showed a consistently high correlation with the crops.

2.2.3 State-level Statistics Apportioned Using Ag Density

The final apportioning procedure analyzed was to use state-level statistics and apportion them to the quadrant. The equation employed in the procedure is listed in Appendix 2. A Spearman correlation was again used and the results are presented in Table 3. The unusual situation in this table is that the coefficients are the same for the four crops for each variable. No variable other than ag density indicated a very high correlation with all of the crops.

2.2.4 County Totals for the Three Procedures

Tables 4-7 list the actual county statistics and the three values for each county for each procedure for each crop. The county total for the procedure employing state-level statistics was determined by summing the values apportioned to each quadrant in the county. The totals at the bottom of each column indicate that very little difference existed between the three procedures. The major problem lies at the county level. Counties with no area for a crop were

apportioned some area of that crop and counties with a large actual area for a crop were invariably underapportioned. The soybean crop is a prime example (Table 5).

Approximately one-half of the counties had soybean area; yet the apportioning procedure using state-level statistics allocated some soybean area to every county. Similarly, counties with large actual areas were allocated much less than the actual area.

TABLE 1

Correlation Coefficients for the Equally Apportioned Procedure

	Sorghum	Soybean	Corn	Wheat
Soybean	.489			
Corn	.105	.572		
Wheat	.690	.670	.383	
APU	-.420	-.246	.062	-.390
Ag Density	-.560	-.445	-.257	-.546
Soil Rank	-.668	-.494	-.227	-.665
Texture	-.268	-.305	-.084	-.321
Depth	-.486	-.266	.018	-.478
Drainage	-.626	-.400	-.209	-.609
Salt	-.249	-.190	-.252	-.217
Slope	-.144	-.019	.210	-.152
Dummy Texture	.230	.002	.108	.151
Dummy Depth	-.338	-.118	.011	-.328
Dummy Drainage	-.485	-.285	-.056	-.488
Dummy Salt	-.249	-.190	-.252	-.217
Dummy Slope	-.345	.294	.335	.321

TABLE 2

Correlation Coefficients for the Apportioned Procedure
Using County Ag Density

	Sorghum	Soybean	Corn	Wheat
Soybean	.508			
Corn	.378	.623		
Wheat	.727	.696	.488	
APU	-.472	-.274	.159	-.402
Ag Density	.750	.523	.623	.619
Soil Rank	-.674	-.533	-.269	-.688
Texture	-.258	-.313	-.107	-.340
Depth	-.507	-.315	-.161	-.485
Drainage	-.638	-.446	-.323	-.624
Salt	-.217	-.193	-.174	-.225
Slope	-.166	-.052	.061	-.148
Dummy Texture	.232	.031	.131	.144
Dummy Depth	-.348	-.173	-.086	-.330
Dummy Drainage	-.523	-.330	-.231	-.496
Dummy Salt	-.217	-.193	-.174	-.225
Dummy Slope	.332	.296	.293	.336

TABLE 3

Correlation Coefficients for the Apportioned Procedure
Using State Ag Density

	Sorghum	Soybean	Corn	Wheat
Soybean	1.000			
Corn	1.000	1.000		
Wheat	1.000	1.000	1.000	
APU	-.493	-.493	-.493	-.493
Ag Density	1.000	1.000	1.000	1.000
Soil Rank	-.556	-.556	-.556	-.556
Texture	-.200	-.200	-.200	-.200
Depth	-.429	-.429	-.429	-.429
Drainage	-.555	-.555	-.555	-.555
Salt	-.123	-.123	-.123	-.123
Slope	-.129	-.129	-.129	-.129
Dummy Texture	.178	.178	.178	.178
Dummy Depth	-.272	-.272	-.272	-.272
Dummy Drainage	-.486	-.486	-.486	-.486
Dummy Salt	-.123	-.123	-.123	-.123
Dummy Slope	.292	.292	.292	.292

TABLE 4

Sorghum

	ACTUAL	EST[1]	QPPW[2]	QPPS[3]
Calamuchita	3,600	3,001	3,600	3,600
Capital	1,000	2,348	1,000	1,000
Colon	5,700	12,003	5,700	5,700
Cruz Del Eje	1,000	14,938	1,001	1,001
General Roca	68,000	67,451	68,000	68,000
General San Martin	21,800	30,529	21,800	21,800
Ischilin	1,500	5,153	1,500	1,500
Juarez Celman	47,000	41,749	47,000	47,000
Marcos Juarez	32,000	64,711	32,001	32,001
Minas	0	5,153	0	0
Pocho	1,700	5,349	1,700	1,700
Pres. Roque Saenz Pena	60,000	49,642	60,000	60,000
Punilla	1,000	11,546	1,000	1,000
Rio Cuarto	100,000	91,522	100,001	100,001
Rio Primero	23,000	32,682	23,000	23,000
Rio Seco	1,000	29,290	1,001	1,001
Rio Segundo	57,000	33,465	57,000	57,000
San Alberto	1,000	11,938	1,000	1,000
San Javier	0	4,349	0	0
San Justo	172,000	103,068	171,999	171,999
Santa Maria	8,000	18,787	8,000	8,000
Sobremonte	300	391	300	300
Tercero Arriba	39,000	21,723	39,000	39,000
Totoral	4,100	9,067	4,100	4,100
Tulumba	3,300	10,763	3,300	3,300
Union	84,000	55,383	84,00	84,000
TOTAL	737,000	6,001	737,003	737,003

1 State-level statistics apportioned using ag density
2 County-level statistics apportioned using ag density
3 County-level statistics apportioned equally

TABLE 5

Soybean

	ACTUAL	EST[1]	QPPW[2]	QPPS[3]
Calamuchita	2,400	1,059	2,400	2,400
Capital	0	828	0	0
Colon	400	4,234	400	400
Cruz Del Eje	0	5,270	0	0
General Roca	0	23,795	0	0
General San Martin	4,200	10,771	4,200	4,200
Ischilin	0	1,818	0	0
Juarez Celman	16,000	14,728	16,000	16,000
Marcos Juarez	82,000	22,829	82,001	82,001
Minas	0	1,818	0	0
Pocho	0	1,887	0	0
Pres. Roque Saenz Pena	0	17,513	0	0
Ponilla	0	4,073	0	0
Rio Cuarto	6,000	32,287	5,998	5,999
Rio Primero	3,000	11,529	3,000	3,000
Rio Seco	0	10,333	0	0
Rio Segundo	29,000	11,806	29,000	29,000
San Alberto	0	4,211	0	0
San Javier	0	1,887	0	0
San Justo	2,000	36,360	1,998	1,998
Santa Maria	22,000	6,628	22,000	22,000
Sobremonte	0	138	0	0
Tercero Arriba	61,000	7,663	61,000	61,000
Totoral	0	3,199	0	0
Tulumba	0	3,797	0	0
Union	32,000	19,538	32,000	32,000
TOTAL	260,000	259,999	259,997	259,998

1 State-level statistics apportioned using ag density
2 County-level statistics apportioned using ag density
3 County-level statistics apportioned equally

TABLE 6

Corn

	ACTUAL	EST[1]	QPPW[2]	QPPS[3]
Calamuchita	8,500	2,687	8,500	8,500
Capital	8,500	2,687	2,000	2,000
Colon	6,500	10,749	6,500	6,500
Cruz Del Eje	1,700	13,378	1,699	1,699
General Roca	11,500	60,403	11,498	11,498
General San Martin	5,000	27,399	4,500	4,500
Ischilin	2,000	4,615	2,000	2,000
Juarez Celman	25,000	37,387	25,000	25,000
Marcos Juarez	174,000	57,950	174,000	174,000
Minas	2,000	4,615	2,000	2,000
Pocho	50,000	4,790	50,001	50,000
Pres. Roque Saenz Pena	7,000	44,455	6,999	6,999
Punilla	2,800	10,340	2,800	2,800
Rio Cuarto	226,000	81,959	226,000	225,999
Rio Primero	4,000	29,267	4,000	4,000
Rio Seco	4,000	26,229	3,999	3,999
Rio Segundo	11,000	29,968	10,999	11,000
San Alberto	14,000	10,690	14,000	14,000
San Javier	2,000	4,790	2,000	2,000
San Justo	4,000	92,299	4,000	4,001
Santa Maria	9,000	16,824	9,000	9,000
Sobremonte	1,000	351	1,000	1,000
Tercero Arriba	17,000	19,453	17,000	17,000
Totoral	7,000	8,120	7,000	7,000
Tulumba	3,000	9,639	3,001	3,001
Union	60,000	49,596	59,999	59,999
TOTAL	660,000	660,640	659,995	659,995

1 State-level statistics apportioned using ag density
2 County-level statistics apportioned using ag density
3 County-level statistics apportioned equally

TABLE 7

Wheat

	ACTUAL	EST[1]	QPPW[2]	QPPS[3]
Calamuchita	3,600	2,036	3,600	3,600
Capital	0	1,593	0	0
Colon	500	8,143	500	500
Cruz Del Eje	0	10,135	0	0
General Roca	22,000	45,760	21,999	22,000
General San Martin	10,500	20,712	10,500	10,500
Ischilin	0	3,496	0	0
Juarez Celman	15,000	28,324	14,999	14,999
Marcos Jurez	180,000	43,902	180,001	180,001
Minas	0	3,496	0	0
Pocho	0	3,629	0	0
Pres. Roque Saenz Pena	27,000	33,679	27,000	27,000
Punilla	0	7,833	0	0
Rio Cuarto	10,000	62,091	10,000	10,000
Rio Primero	6,500	22,172	6,500	6,500
Rio Seco	0	19,871	0	0
Rio Segundo	16,000	22,703	16,000	16,001
San Alberto	0	8,099	0	0
San Javier	0	3,629	0	0
San Justo	22,000	69,924	22,001	22,001
Santa Maria	2,900	12,746	2,900	2,900
Sobremonte	0	266	0	0
Tercero Arriba	21,000	14,737	21,000	21,000
Totoral	0	6,152	0	0
Tulumba	0	7,302	0	0
Union	163,000	37,573	162,999	162,999
TOTAL	500,000	500,003	499,999	500,001

1 State-level statistics apportioned using ag density
2 County-level statistics apportioned using ag density
3 County-level statistics apportioned equally

PART 3.0 CONCLUSIONS AND RECOMMENDATIONS

3.1 CONCLUSIONS

The apportioning procedure using county-level statistics and ag density does a very good job of allocating the area statistics to those quadrants with the higher ag density. However, very few countries throughout the world publish county-level statistics; consequently the need to use state level statistics.

Apportioning of state-level statistics using ag density was very poor. Although the state total was completely apportioned to the quadrants, the individual county totals varied greatly from the actual county areas. Use of this procedure would lead to analysts' conclusions concerning cropping practices at odds with reality. An analyst could conclude that a major shift from one crop to another has occurred within a county while in actuality no such shift occurred. Therefore, because very few countries publish county-level statistics and the apportioning procedure employing state-level statistics was very poor, some other method or procedure must be devised to apportion state, or even some higher hierarchial level, statistics to the quadrant.

3.2 RECOMMENDATIONS

Any apportioning procedure must employ data stored in the FCCAD data base. Because of the low correlation between the soils information and the crops, the first recommendation is that more detailed information about the soils be obtained to employ these variables in an apportioning procedure.

Quadrant level ag density is also required in any procedure. The ranges used in the categories in this study may be too broad, especially in the fourth category (5-40 percent). The second recommendation is that categories should be broken down into five percent ranges. Although the small ranges are highly desirable, they may be unfeasable due to limitations of present technology.

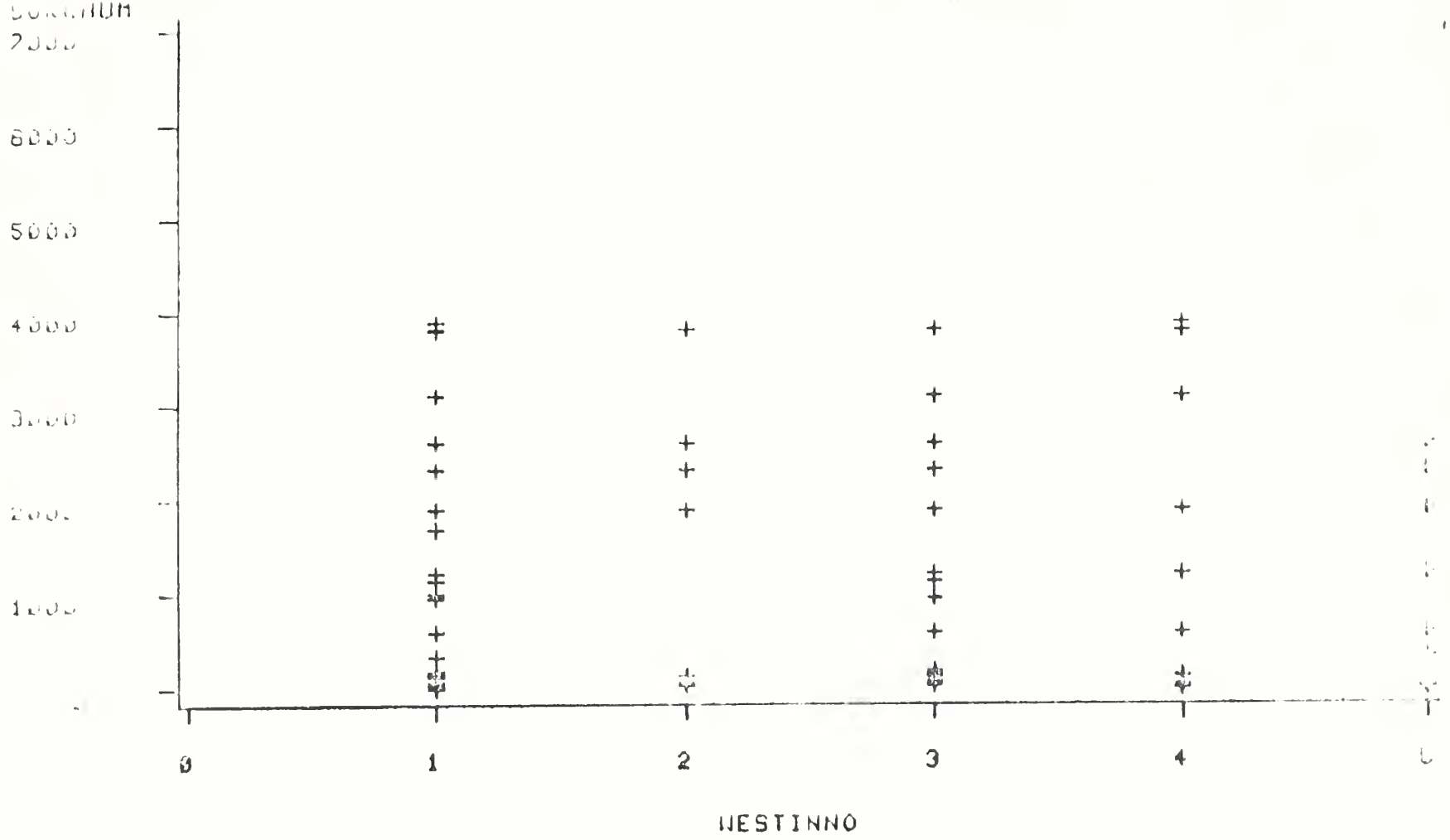
APPENDIX 1

INDEX

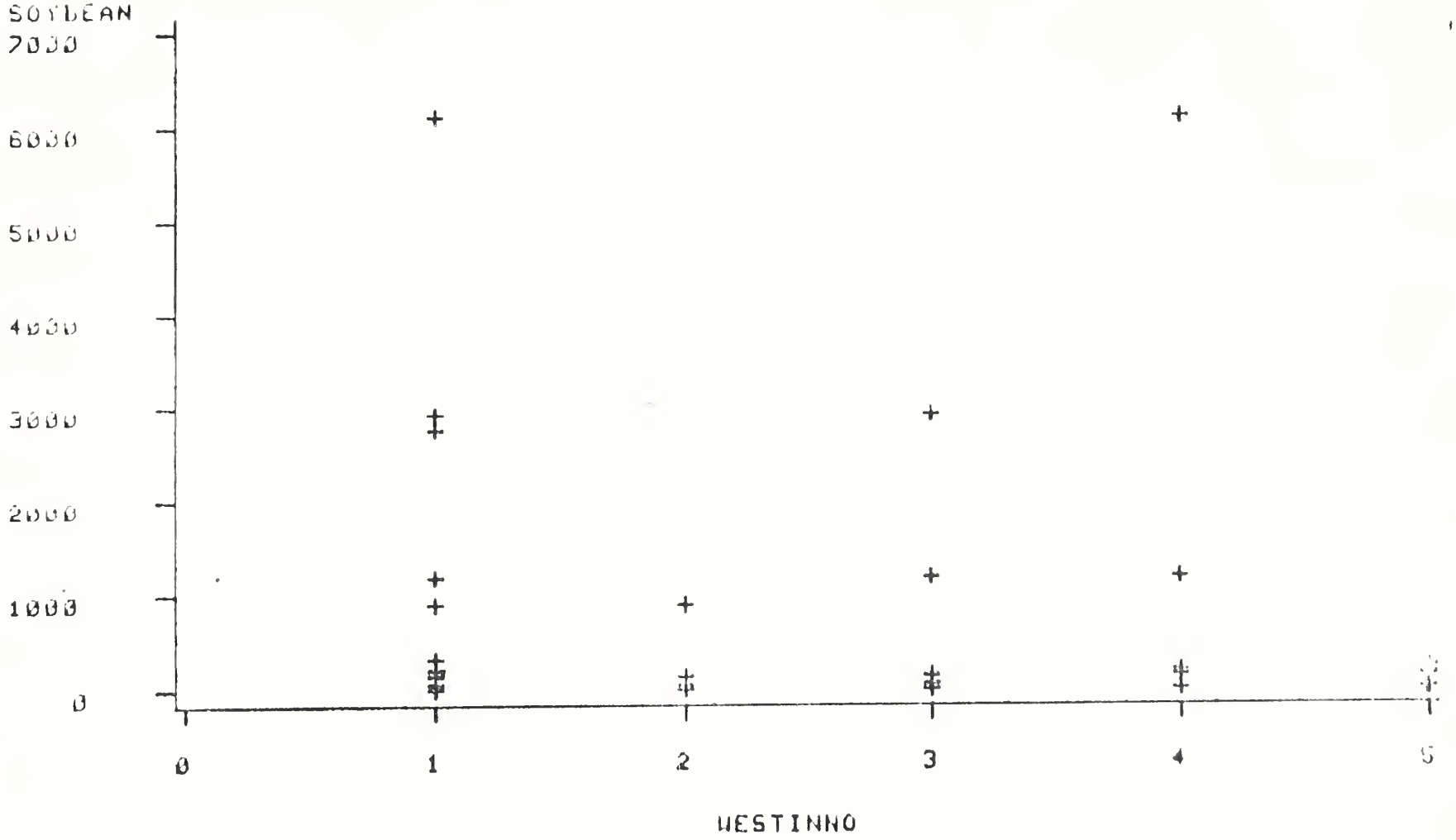
Scatter Plots

SEP 23 1982

SORGHUM VS. AG DENSITY

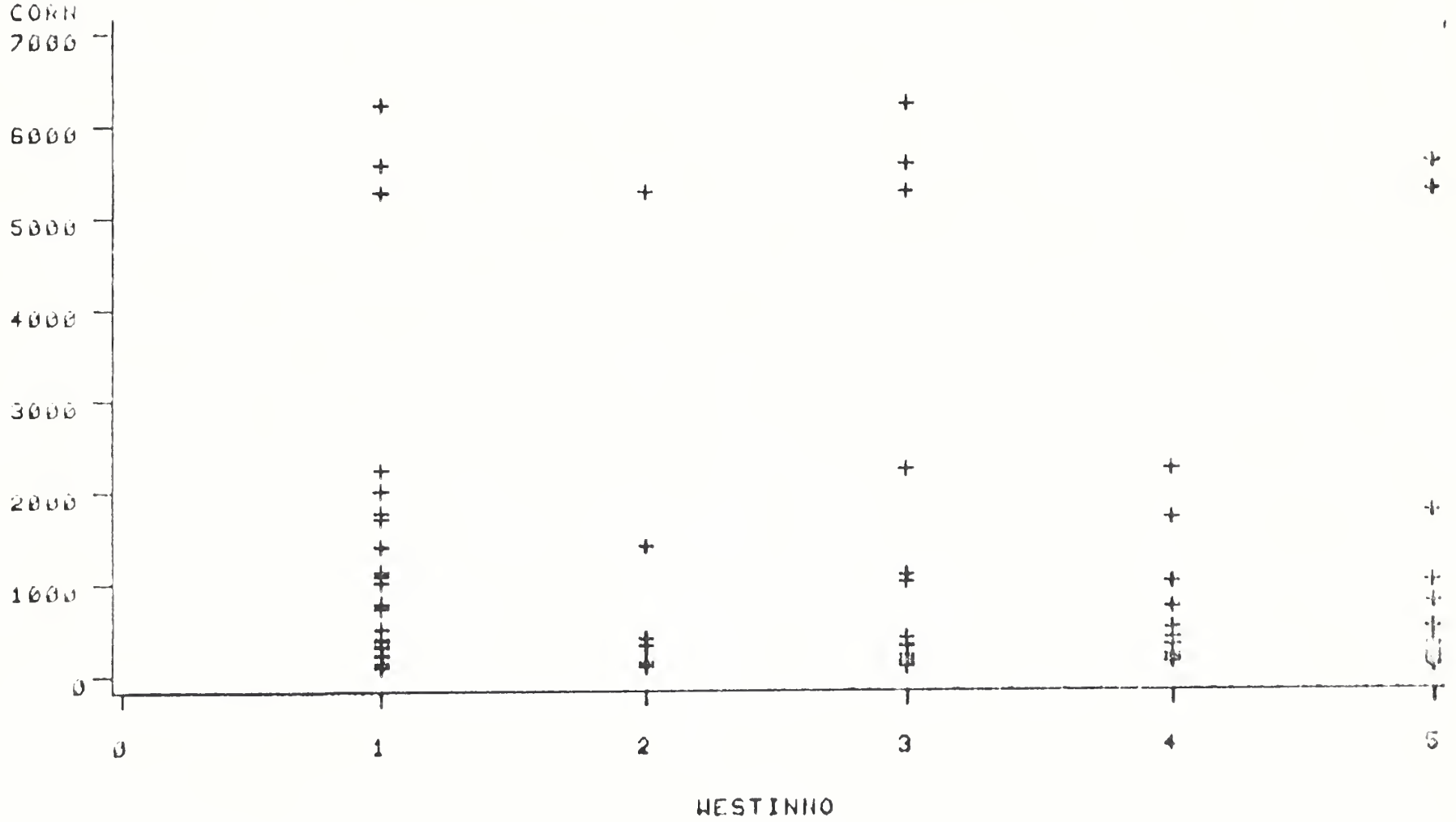


SOYBEAN VS. AG DENSITY

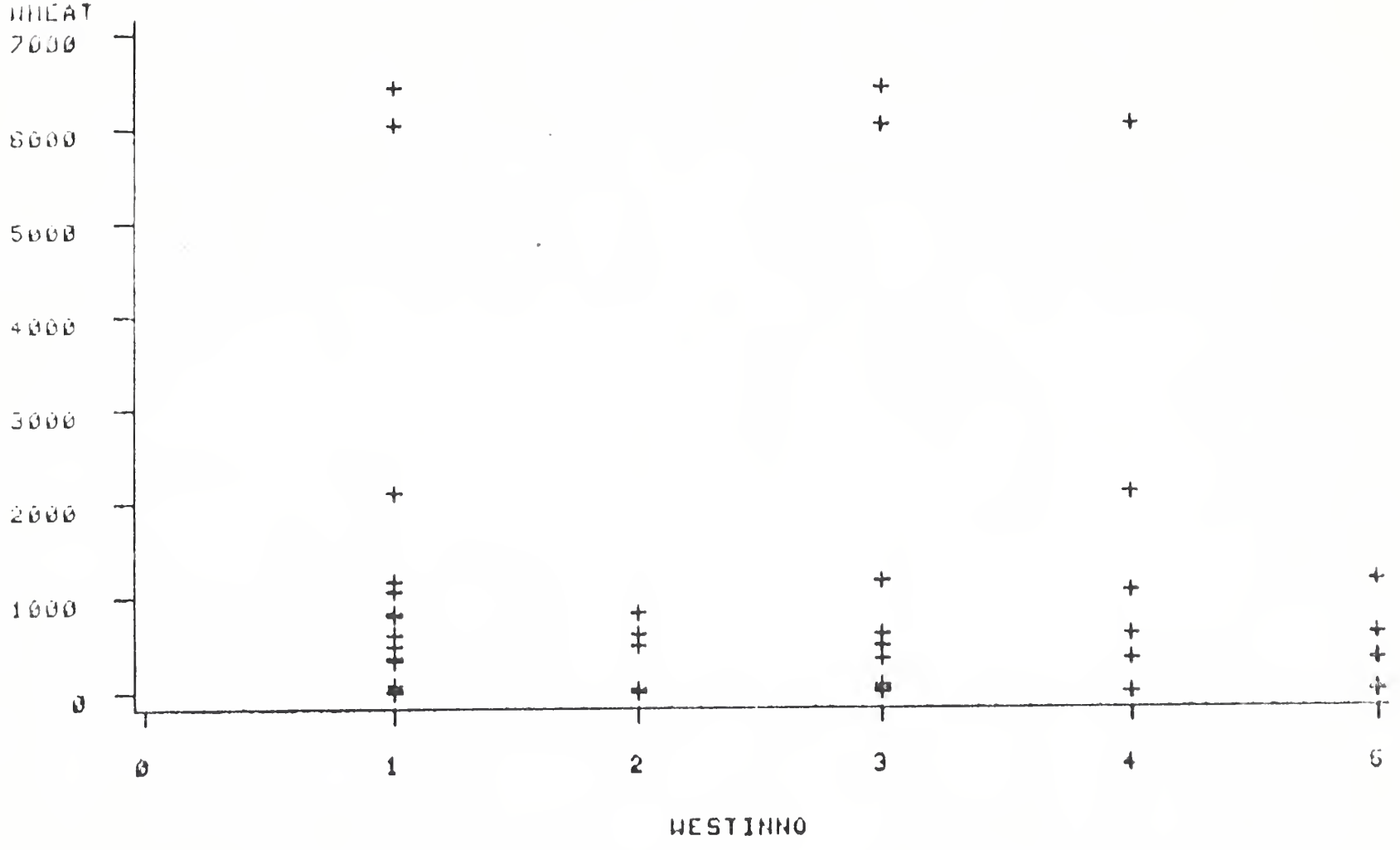


SEP 2 3 1982

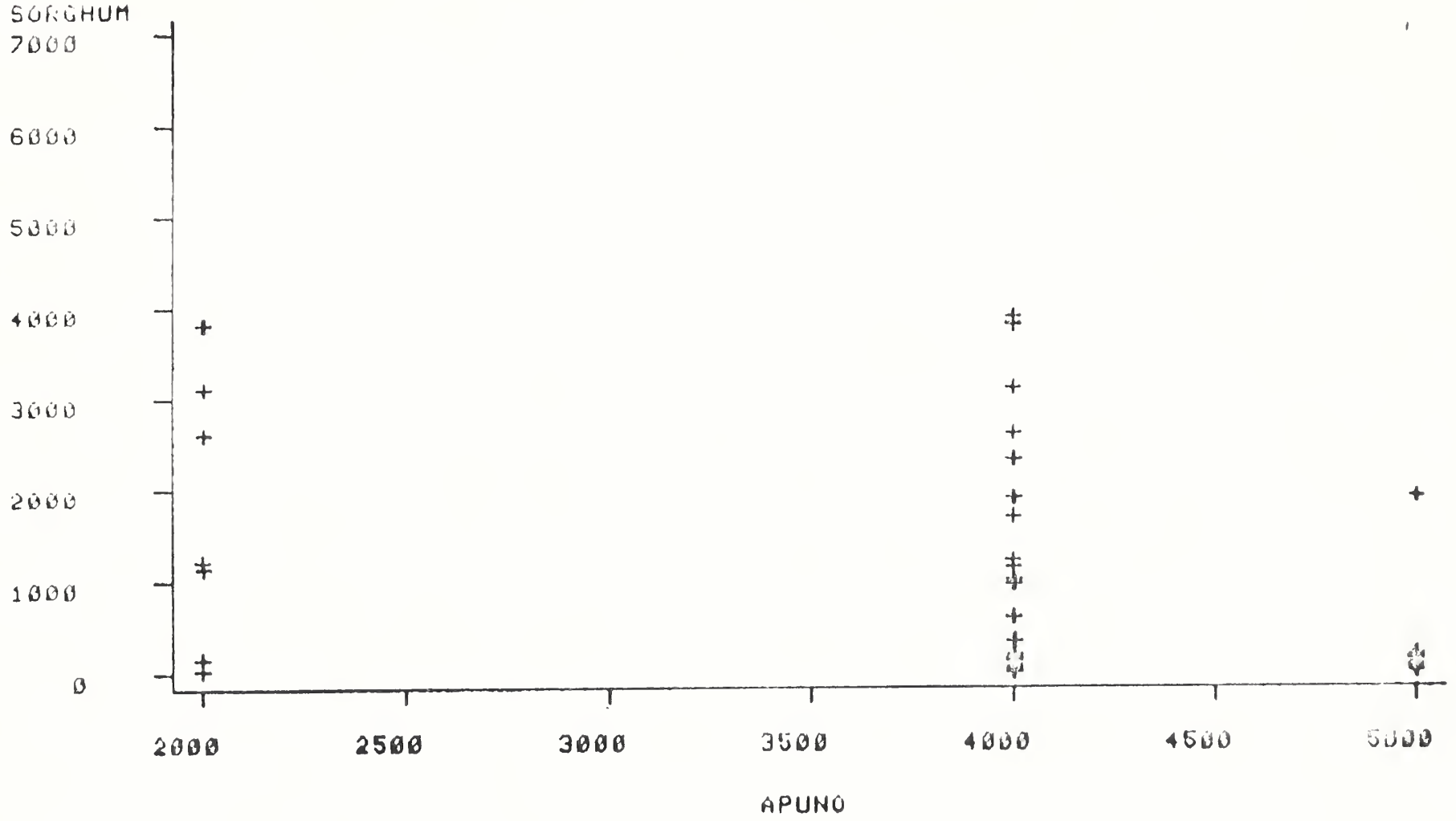
CORN VS. AG DENSITY



WHEAT VS. AG DENSITY

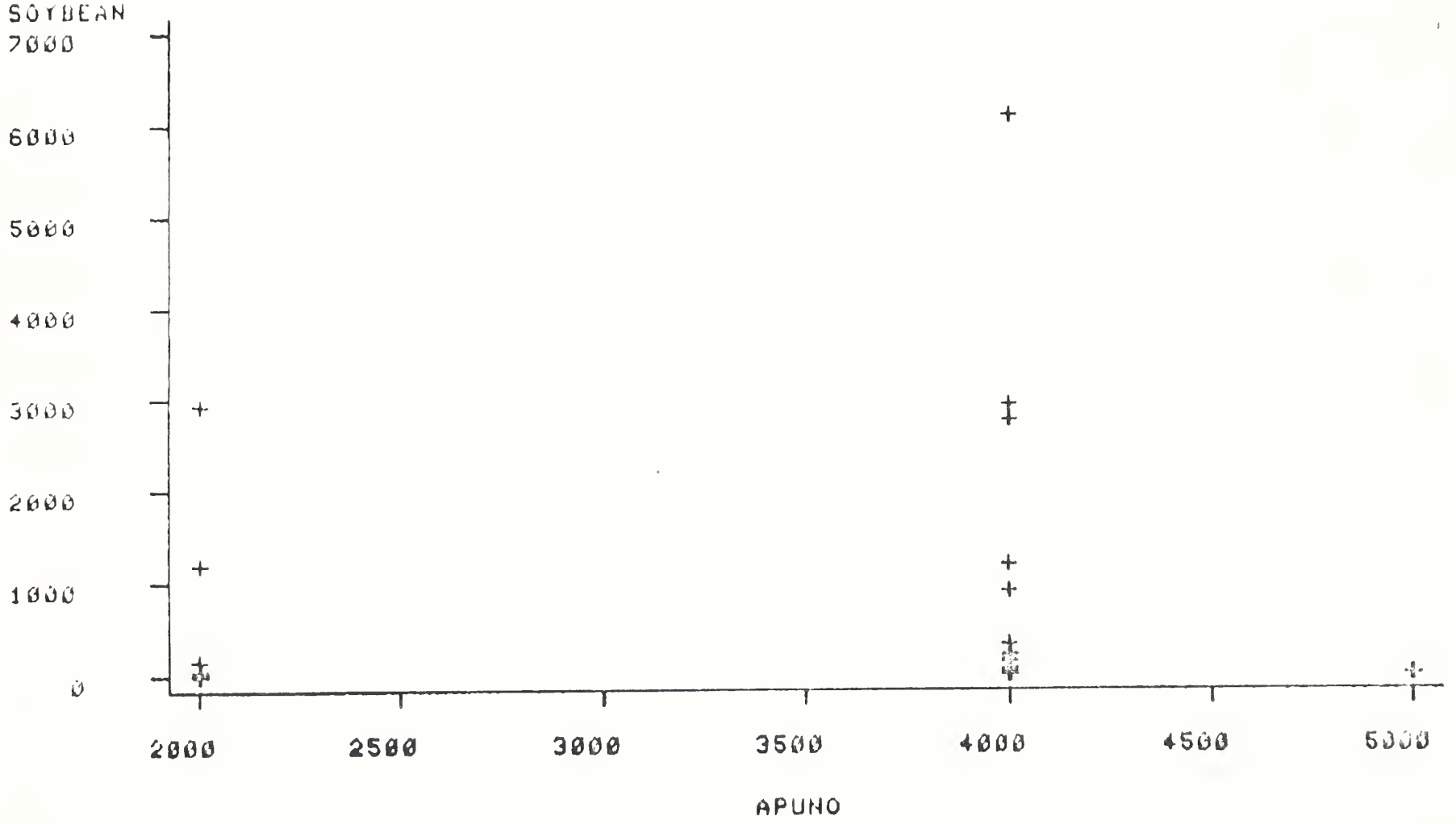


SORGHUM VS. APU

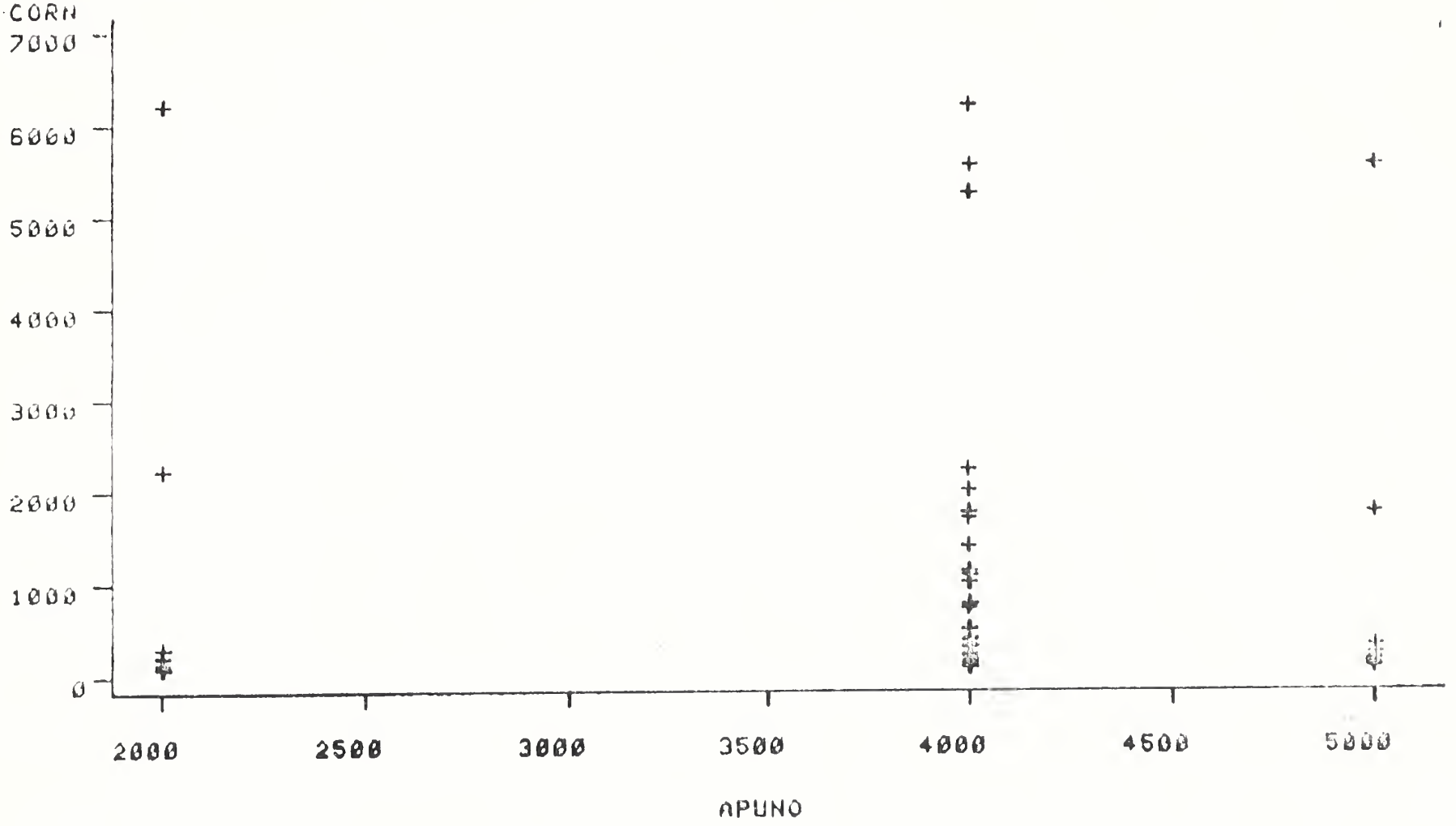


SEP 23 1982

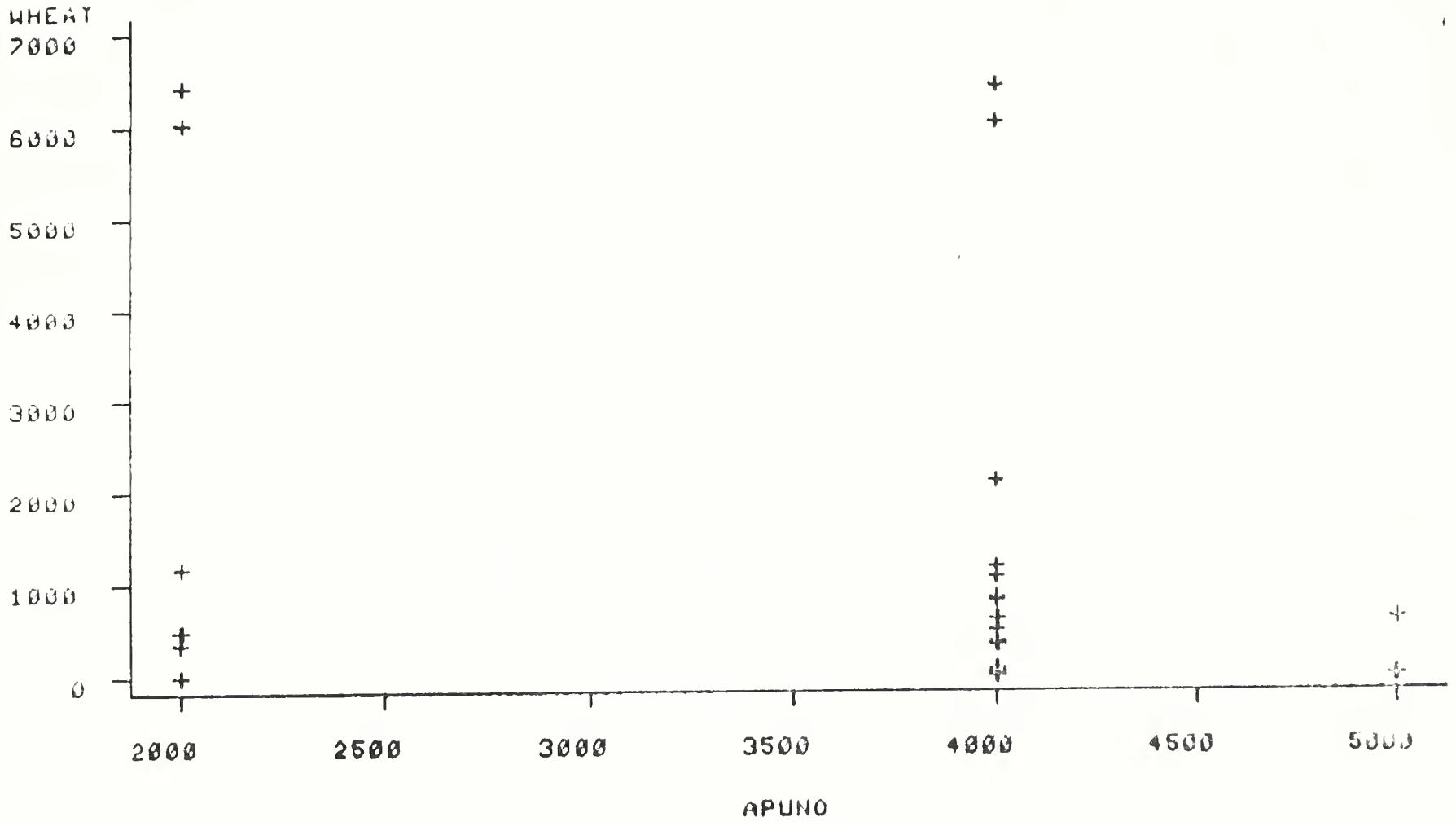
SOYBEAN VS. APU



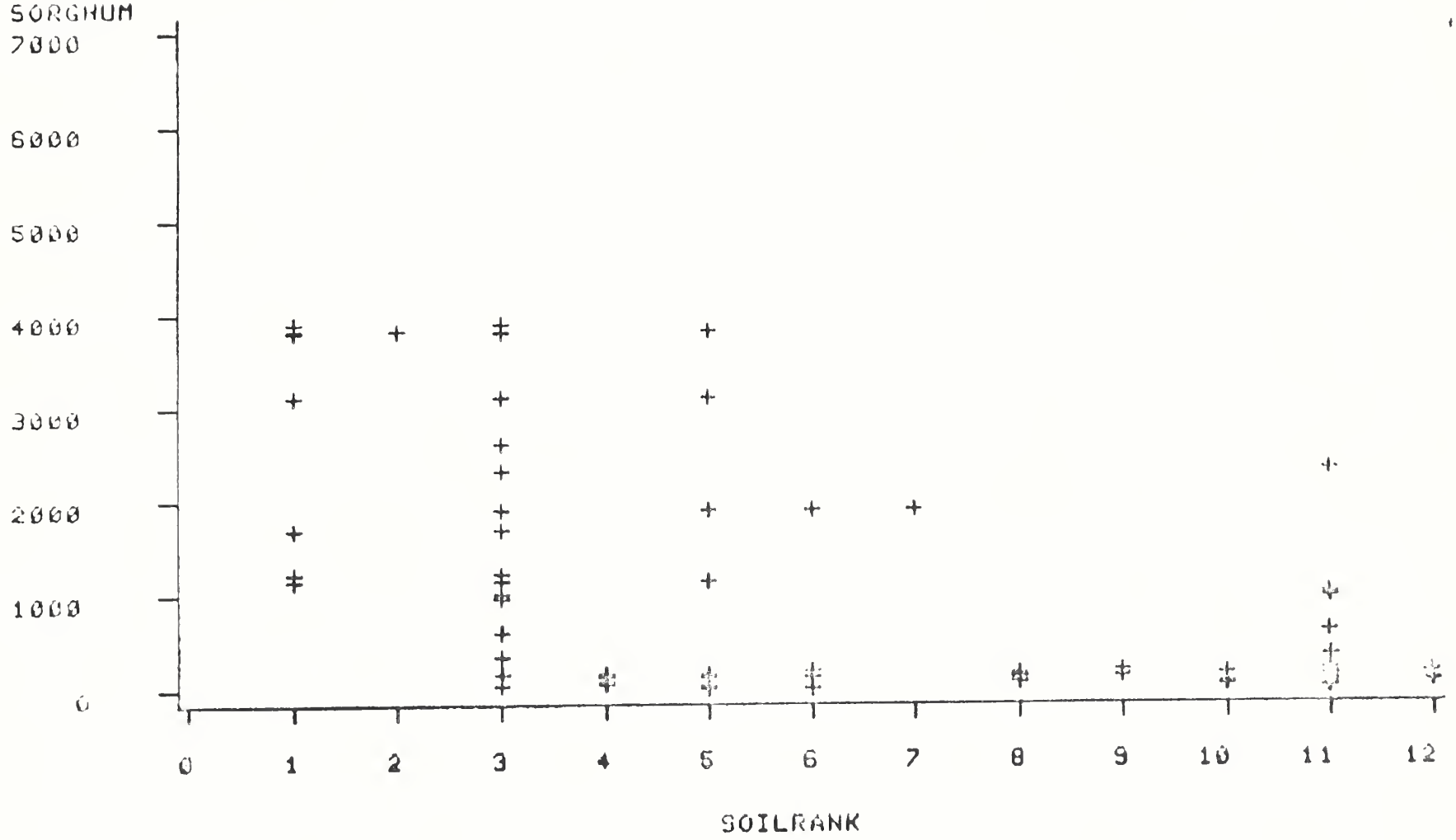
CORN VS. APU


CORN
7000
6000
5000
4000
3000
2000
1000
0
2000
2500
3000
3500
4000
4500
5000
APUNO

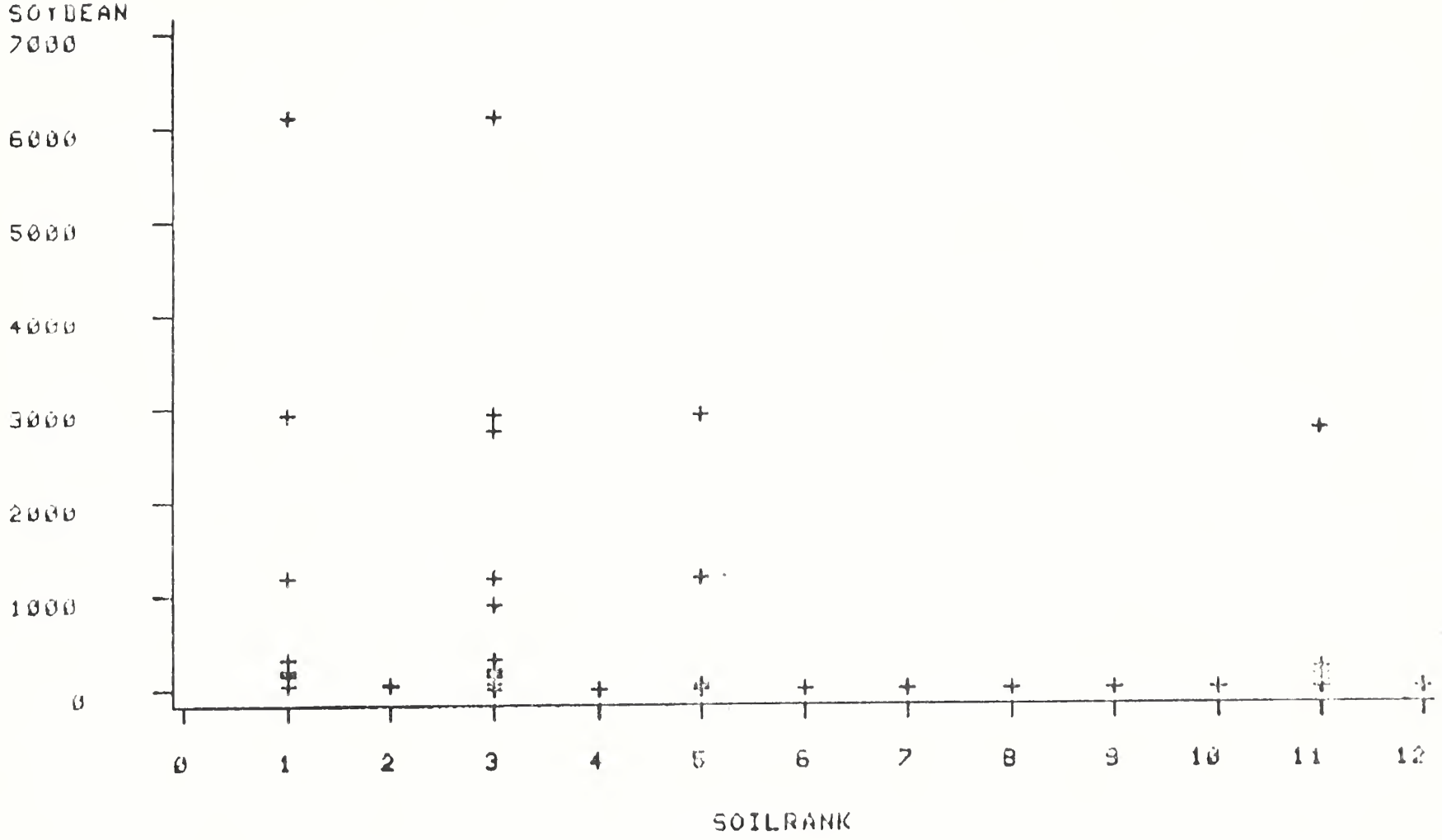
WHEAT VS. APU


WHEAT
7000
6000
5000
4000
3000
2000
1000
0
2000
2500
3000
3500
4000
4500
5000
APUNO

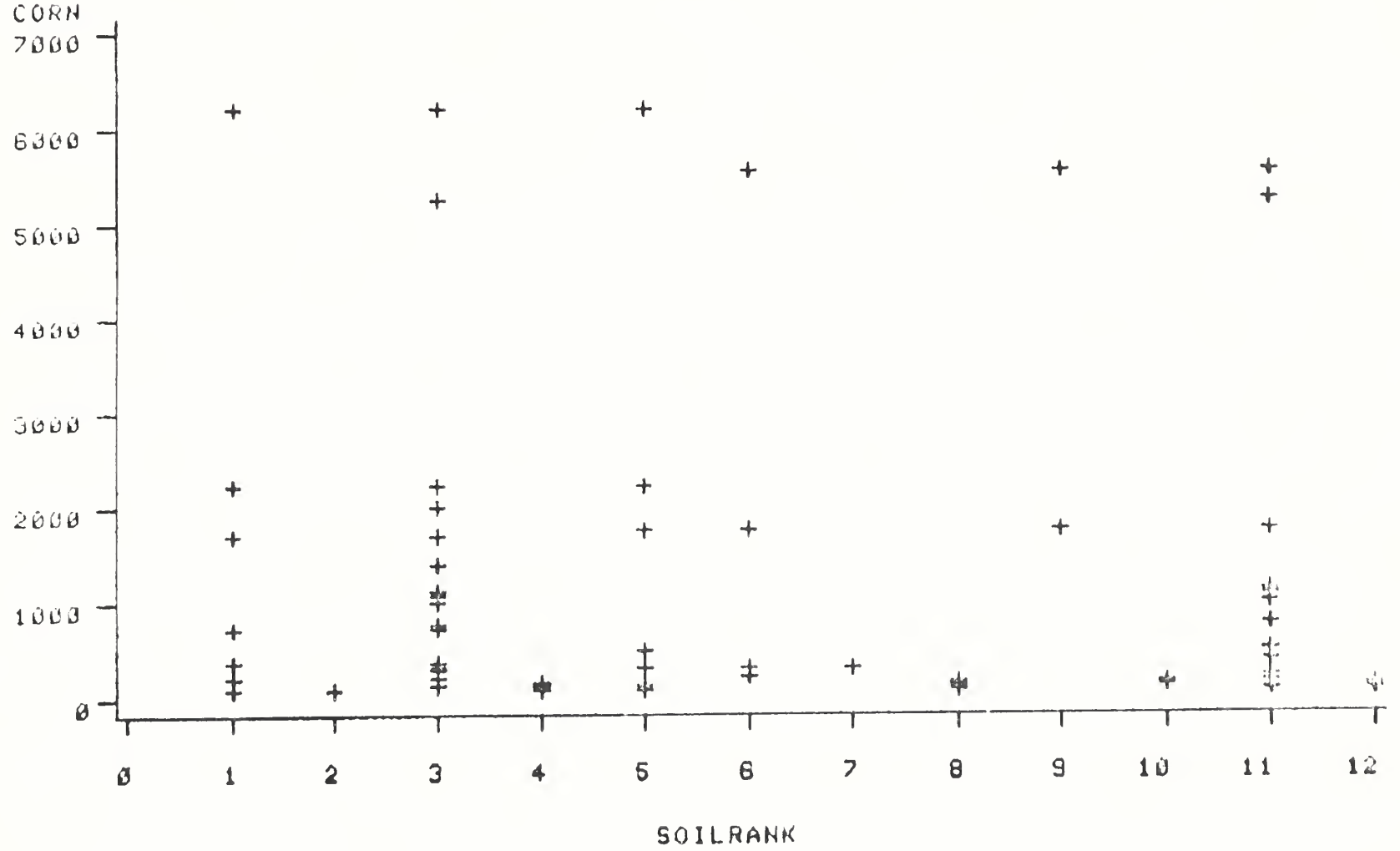
SORGHUM VS. SOIL RANK



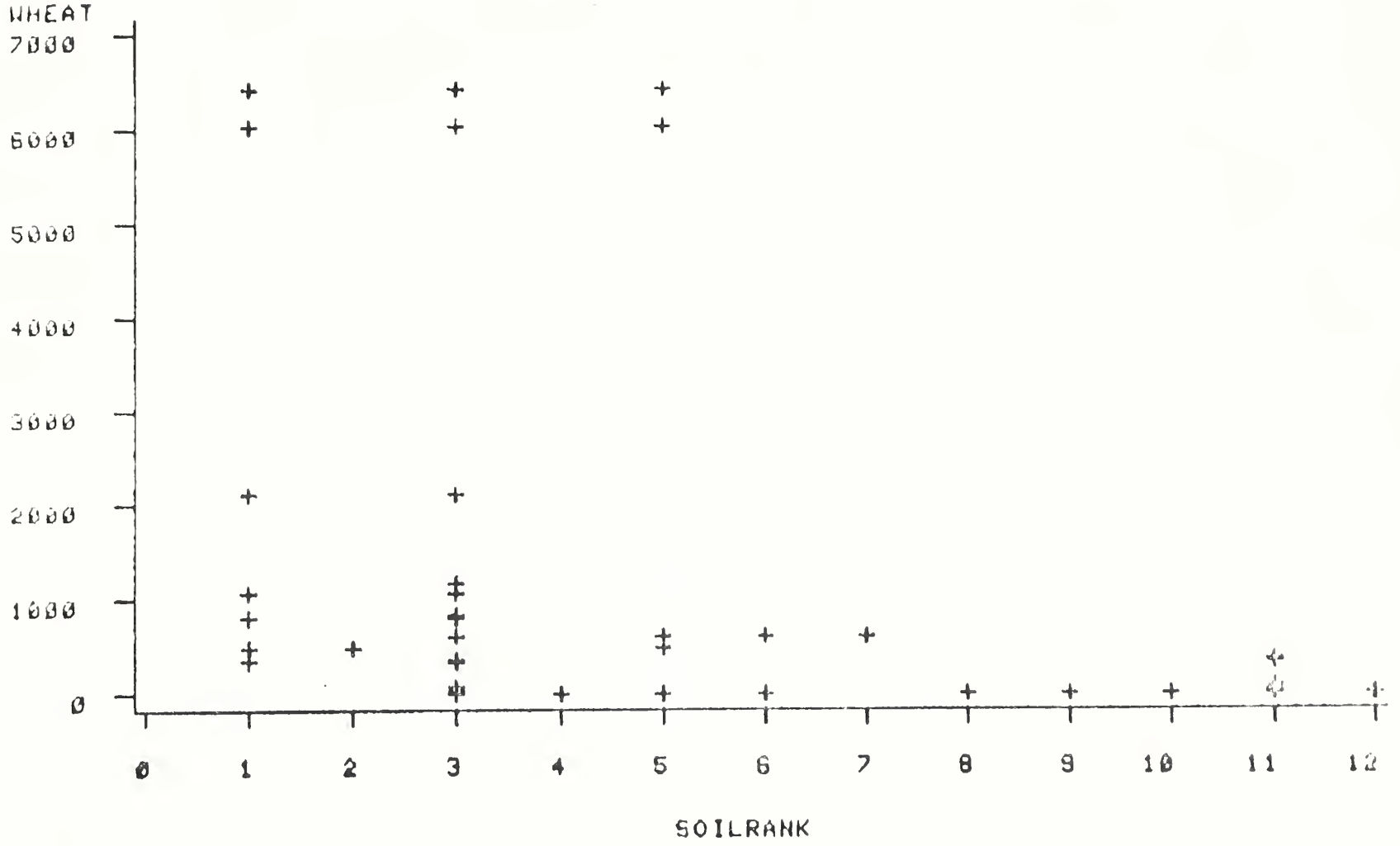
SOYBEAN VS. SOIL RANK



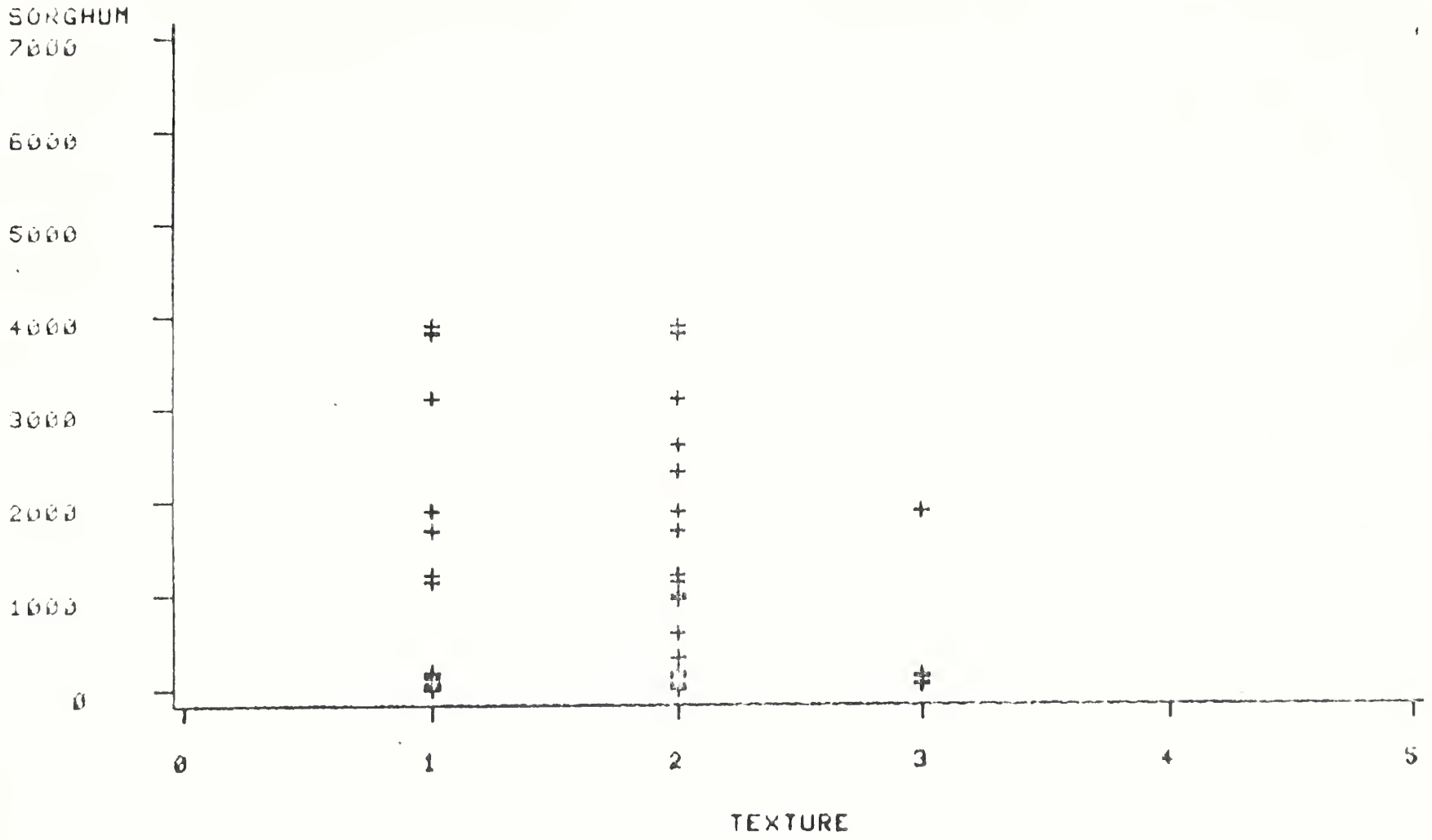
CORN VS. SOIL RANK



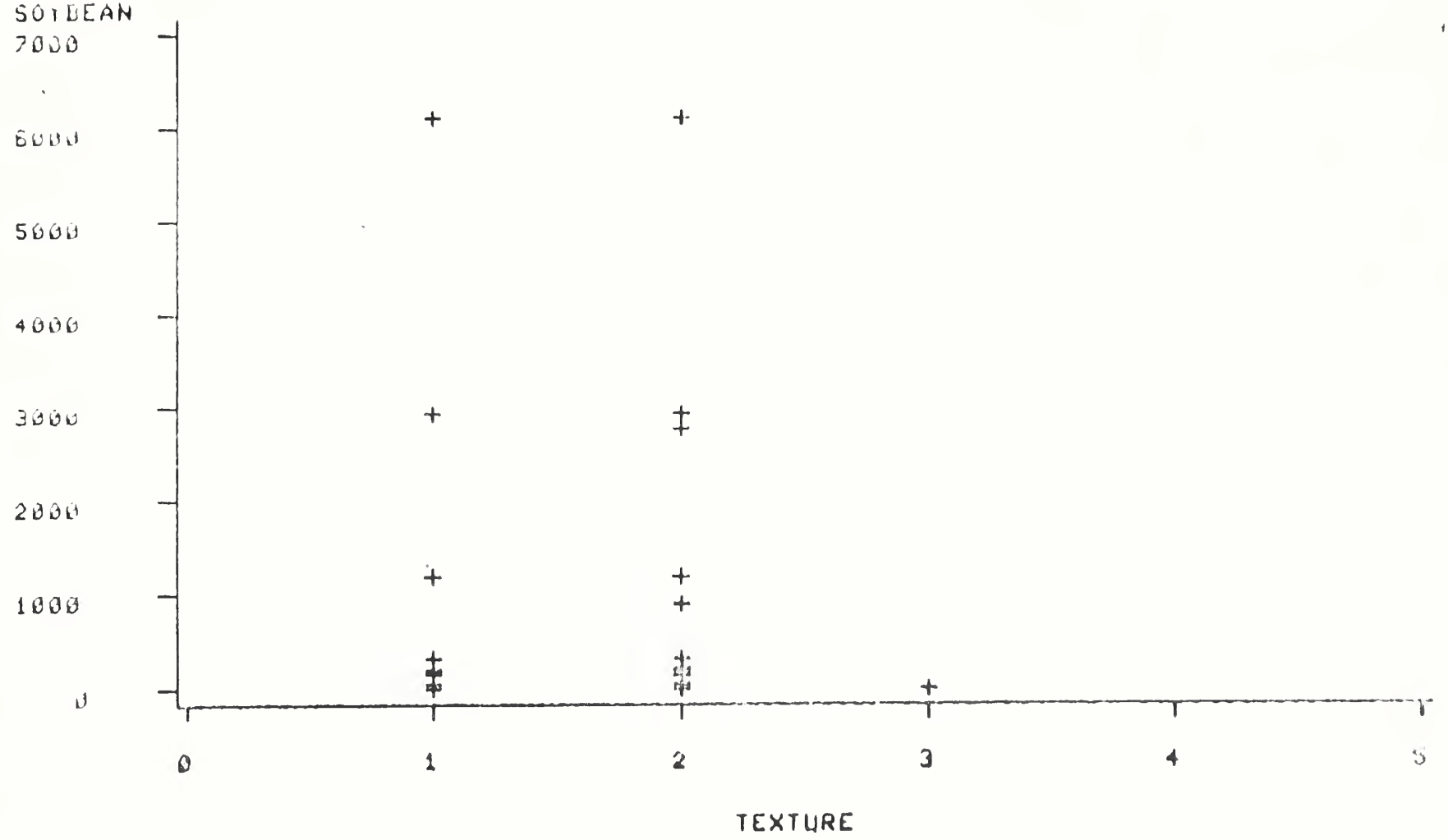
WHEAT VS. SOIL RANK



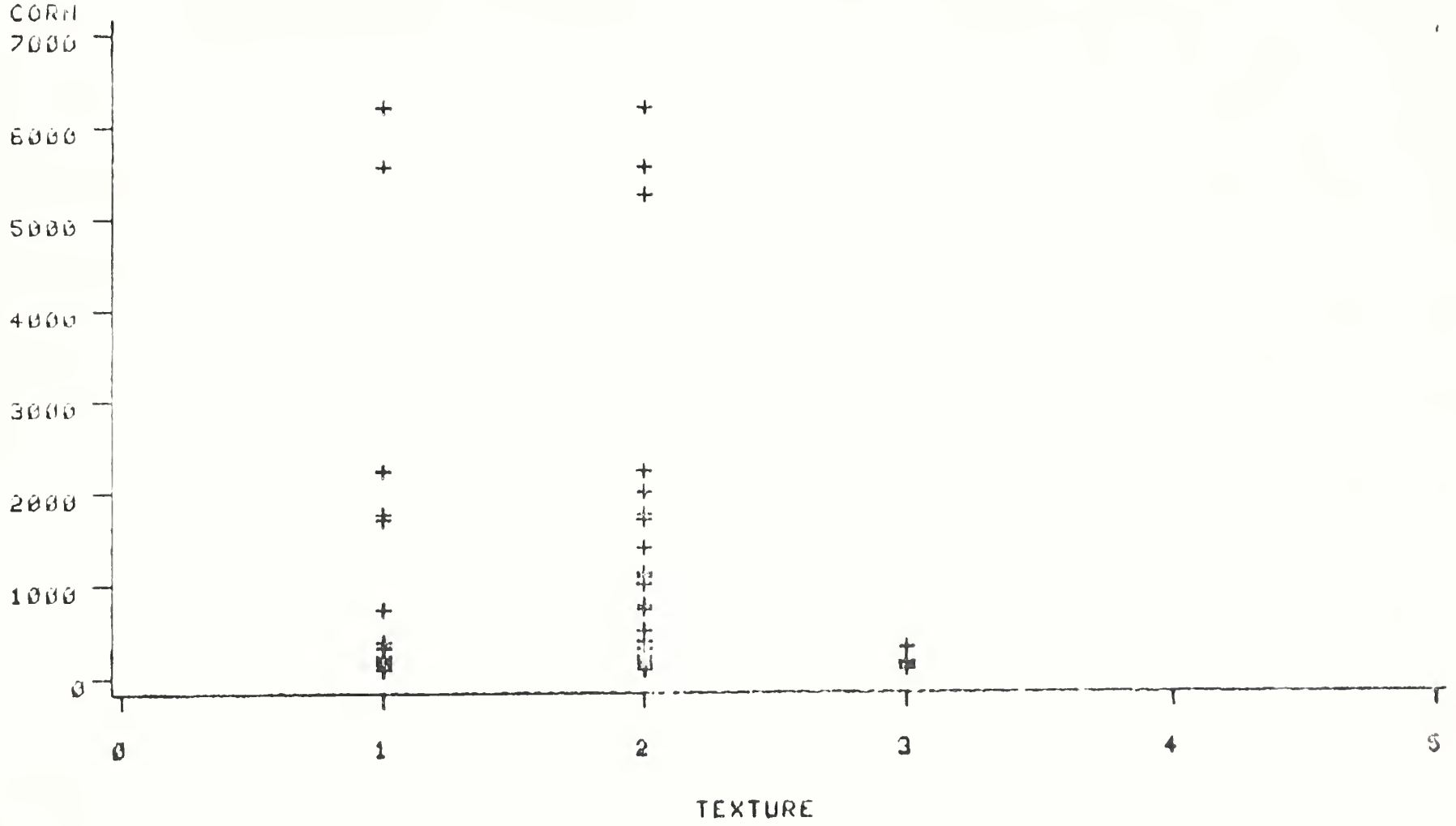
SORGHUM VS. TEXTURE



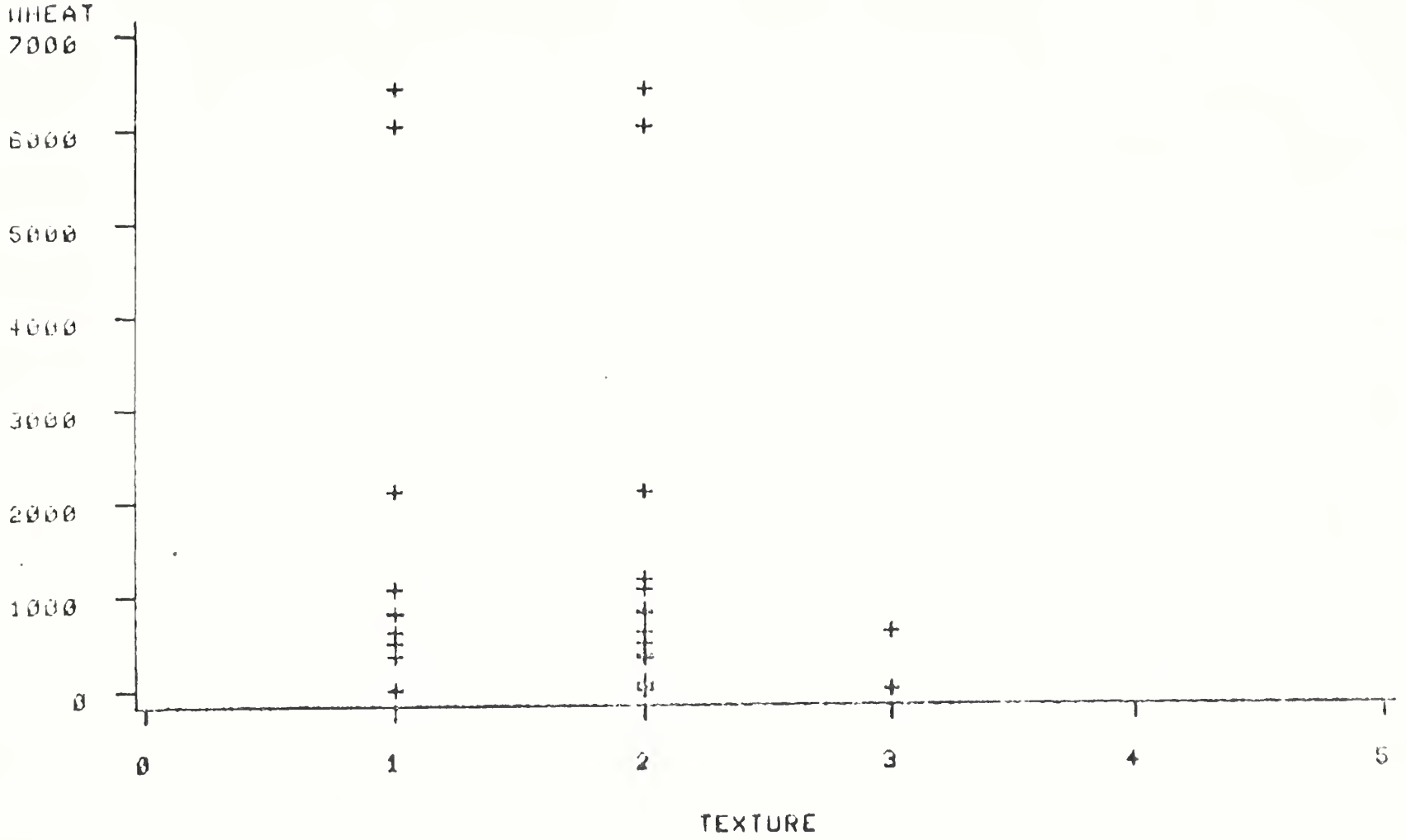
SOYBEAN VS. TEXTURE



CORN VS. TEXTURE

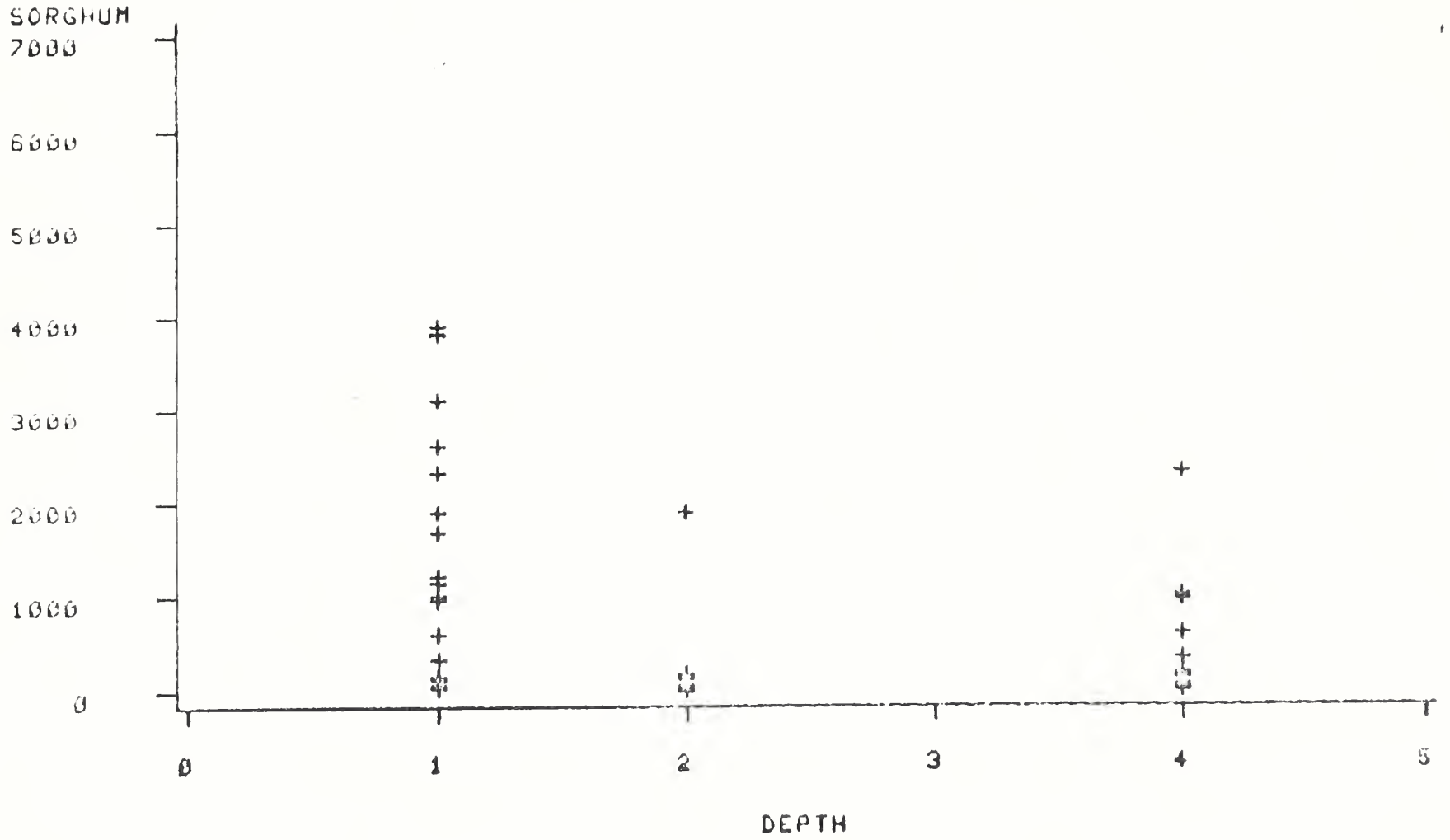


WHEAT VS. TEXTURE

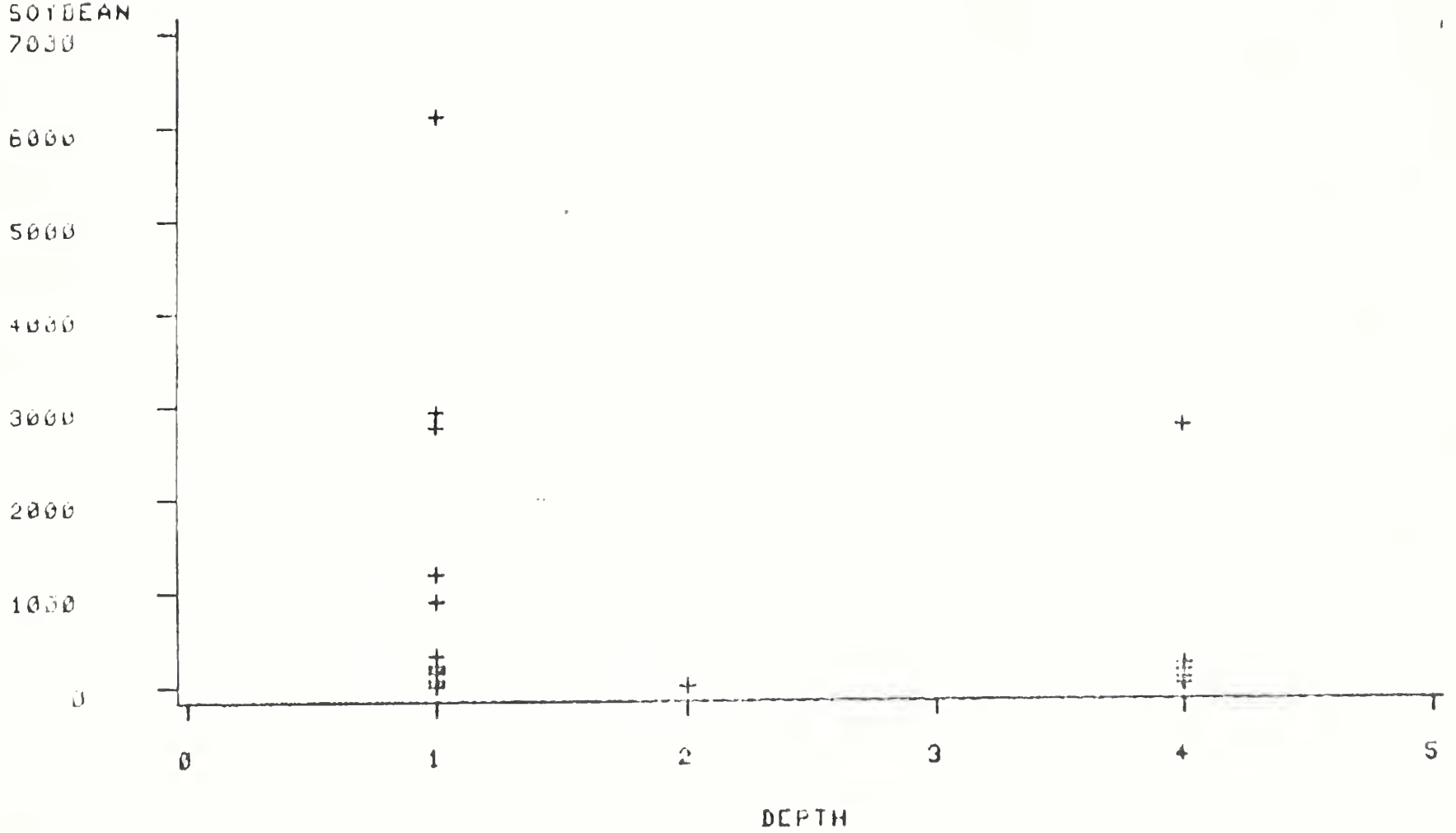


SEP 23 1982

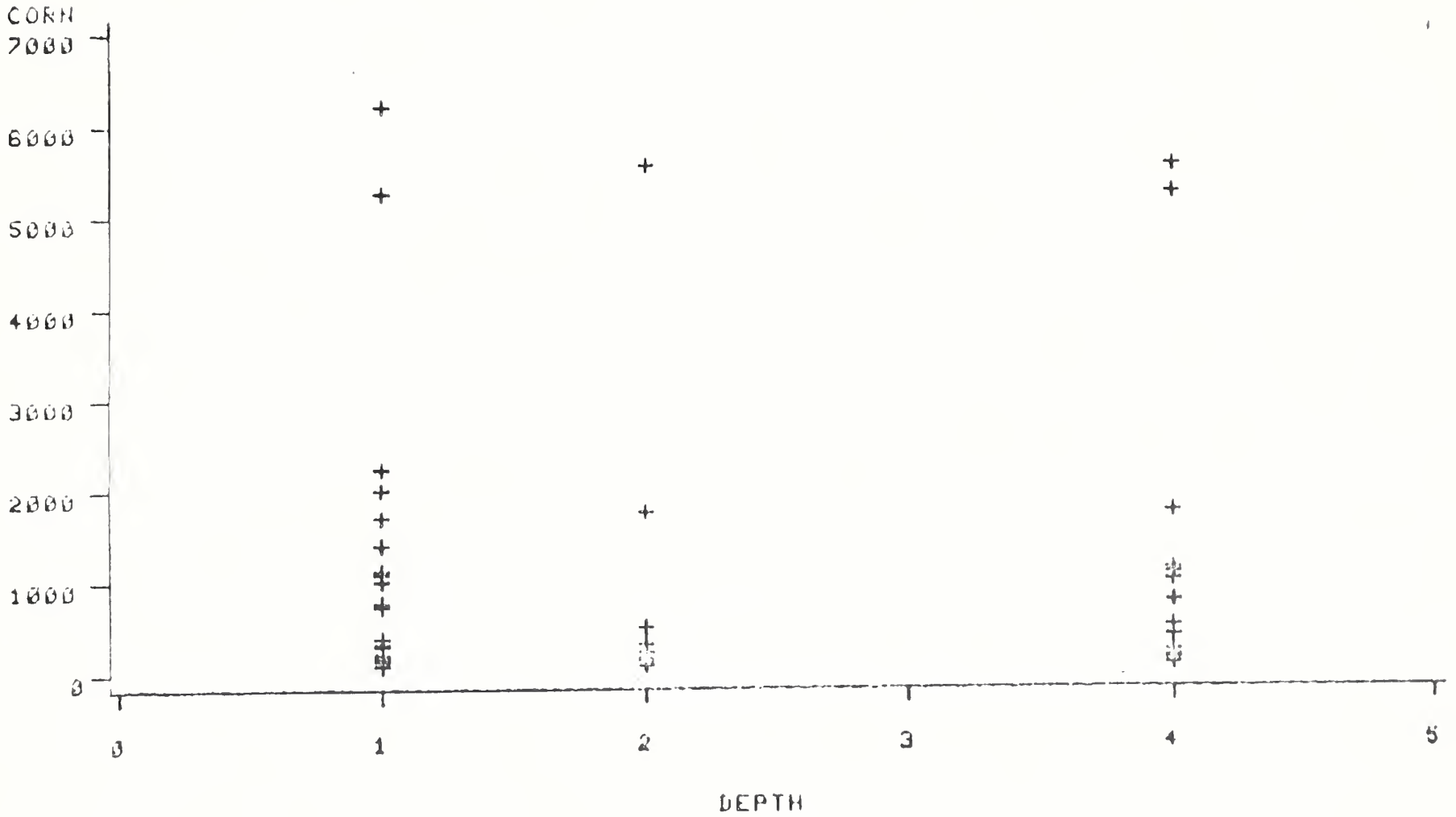
SORGHUM VS. DEPTH



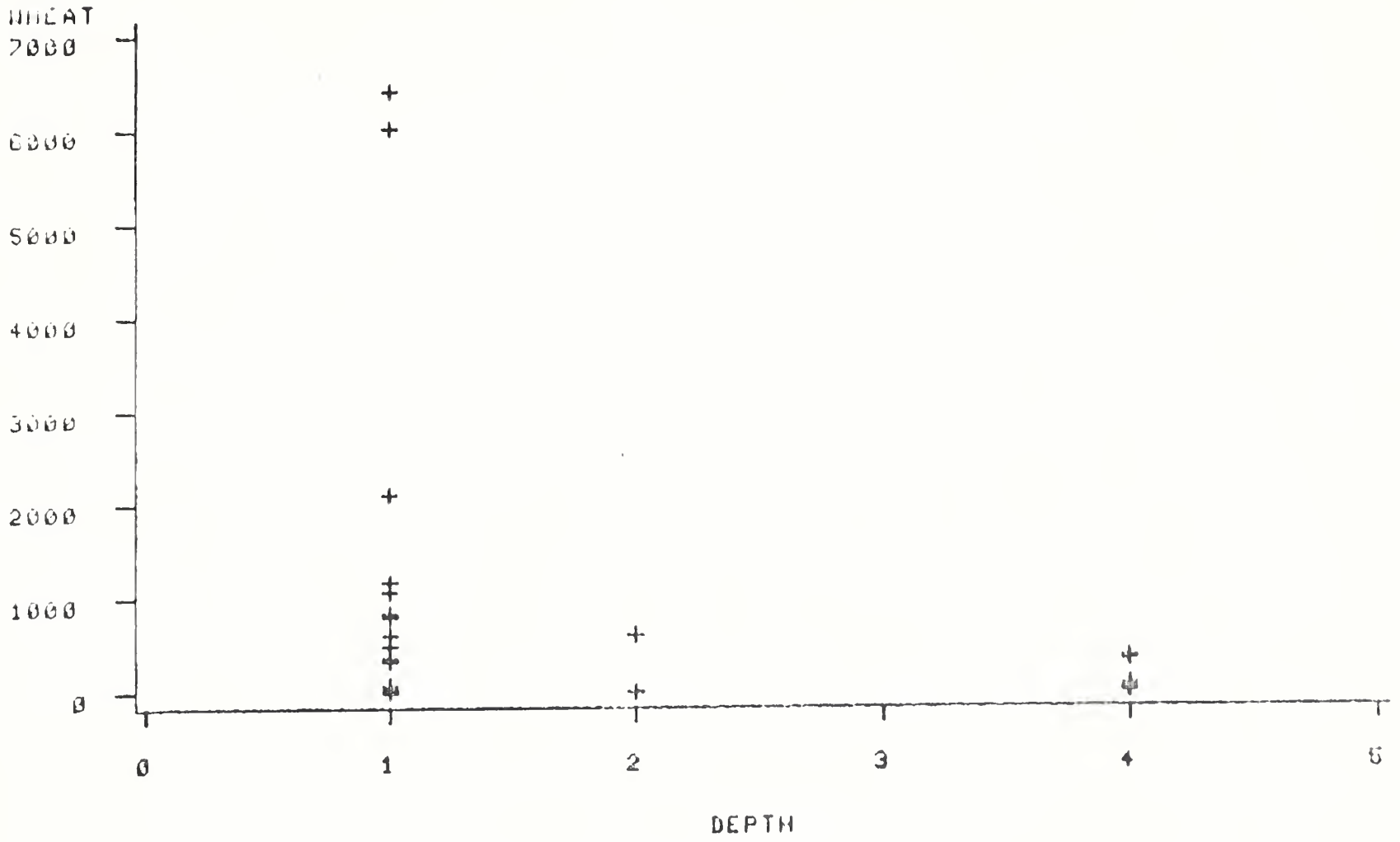
SOYBEAN VS. DEPTH



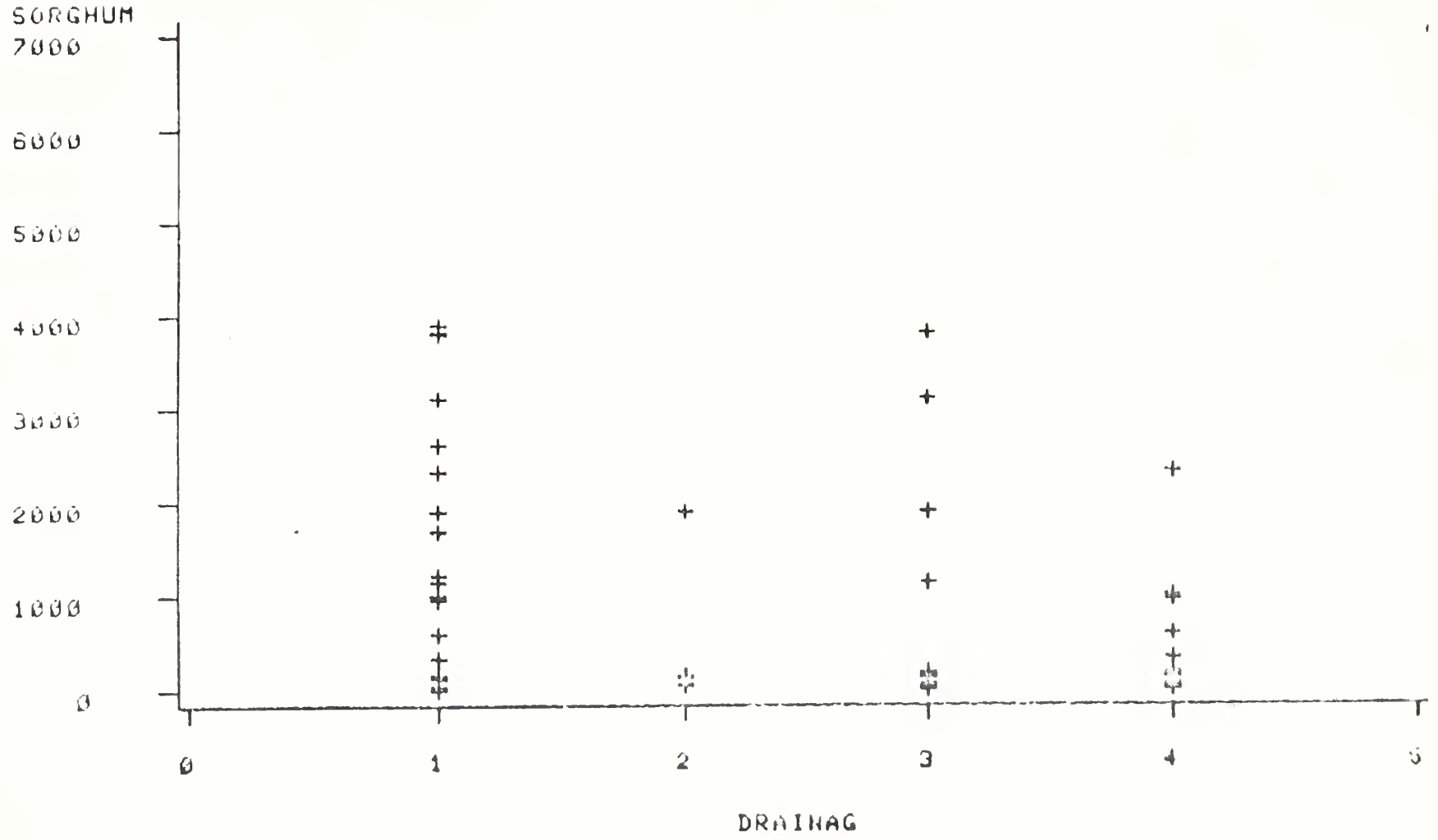
CORN VS. DEPTH



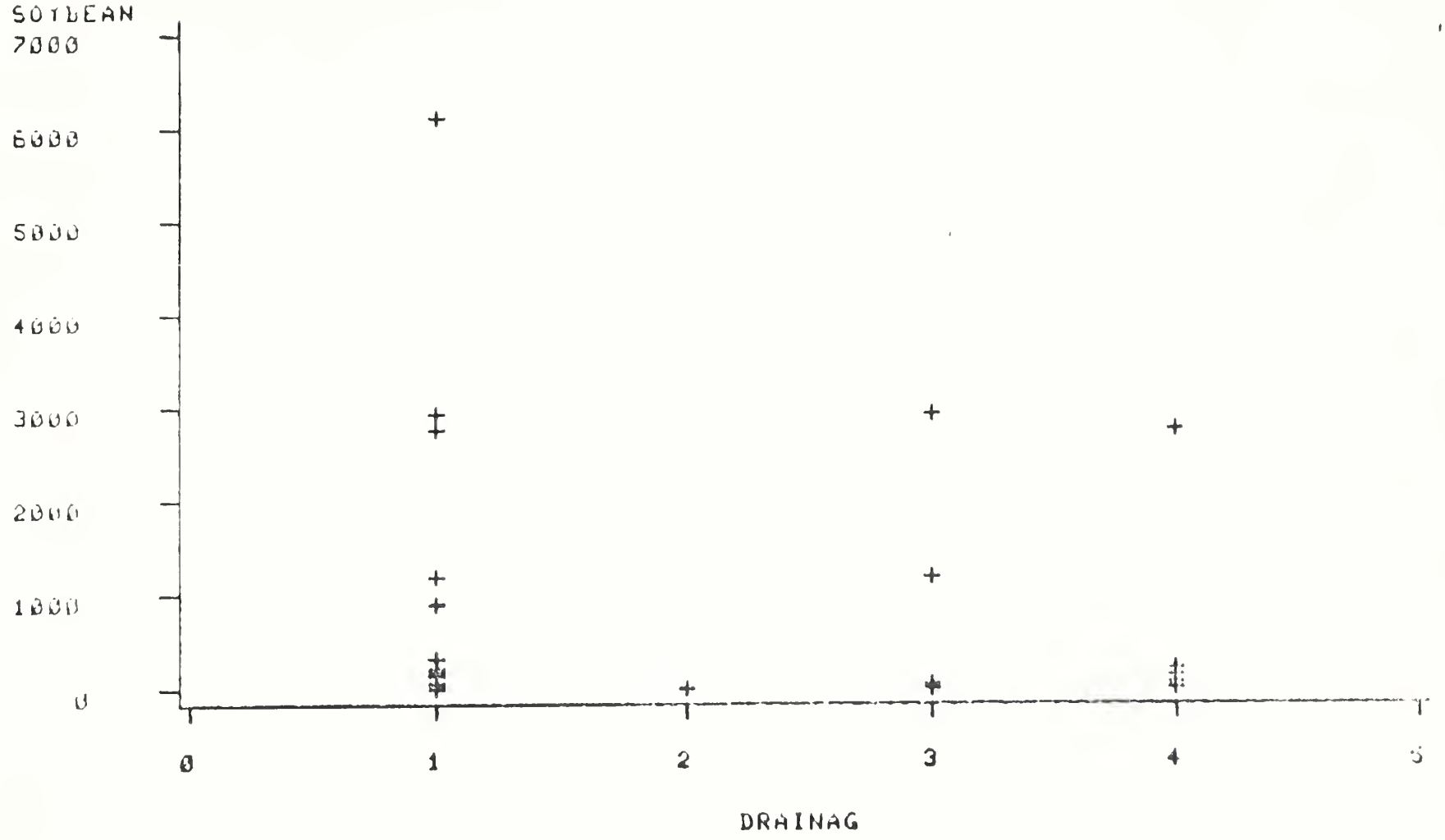
WHEAT VS. DEPTH



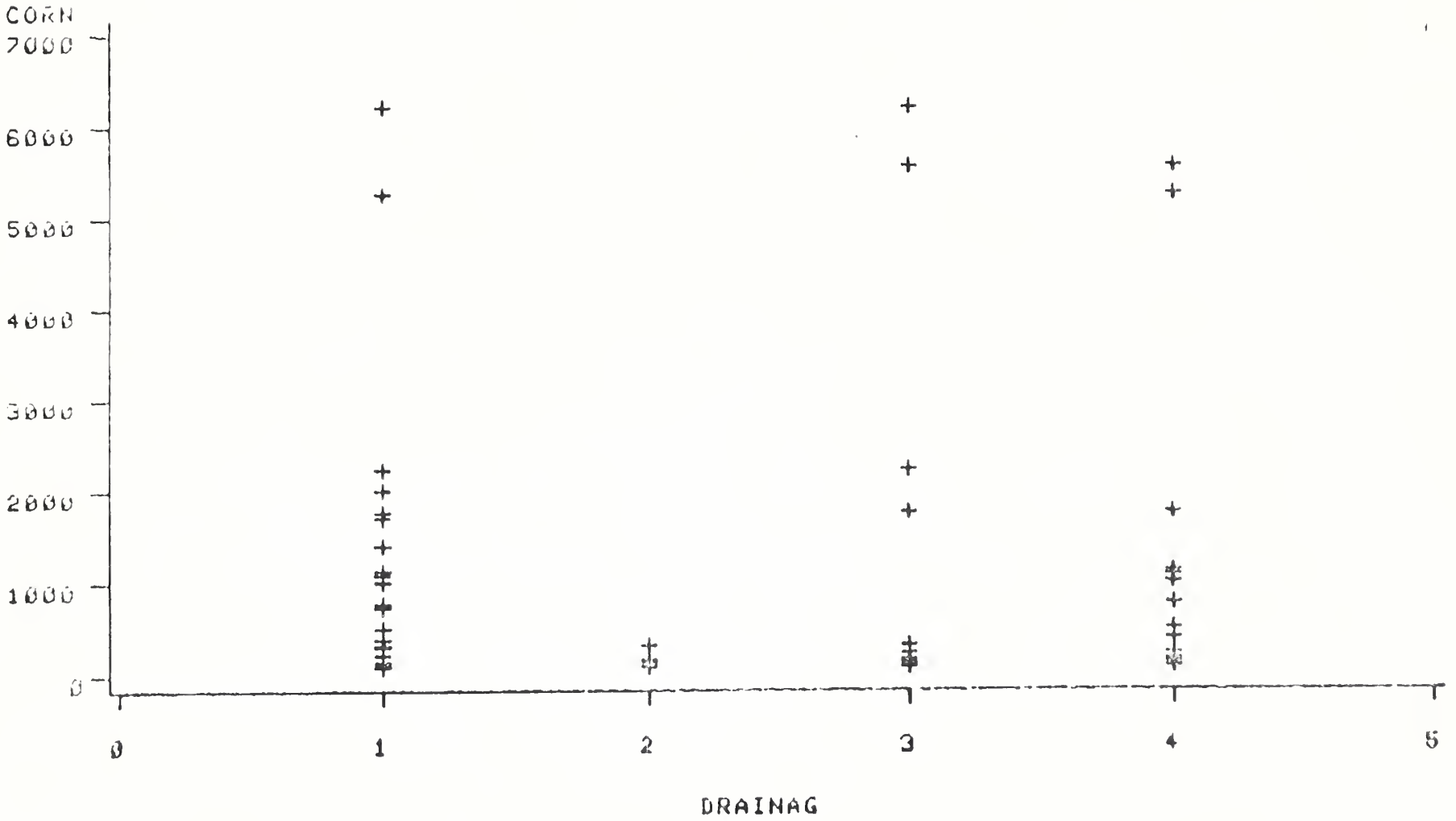
SORGHUM VS. DRAINAGE



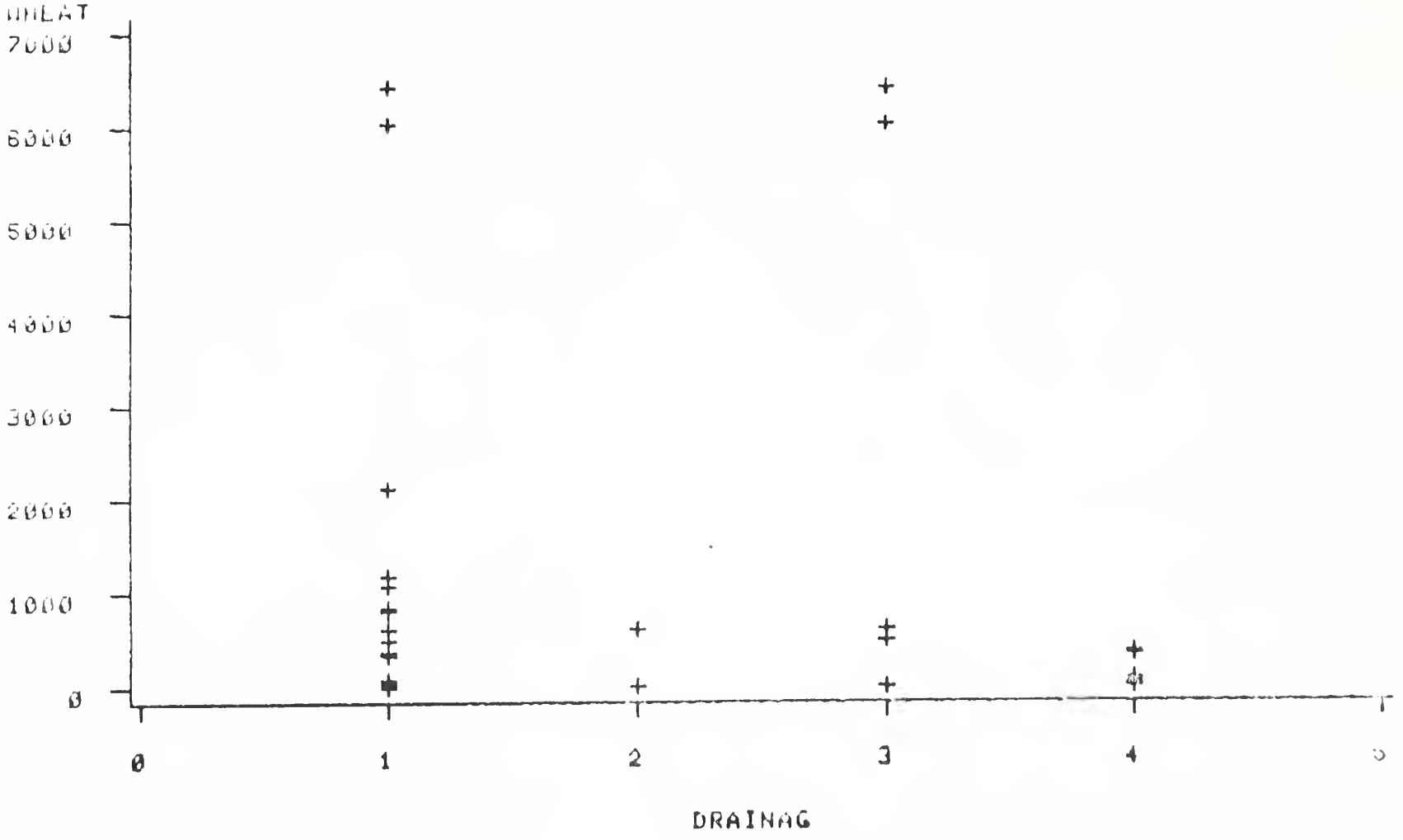
SOYBEAN VS. DRAINAGE



CORN VS. DRAINAGE

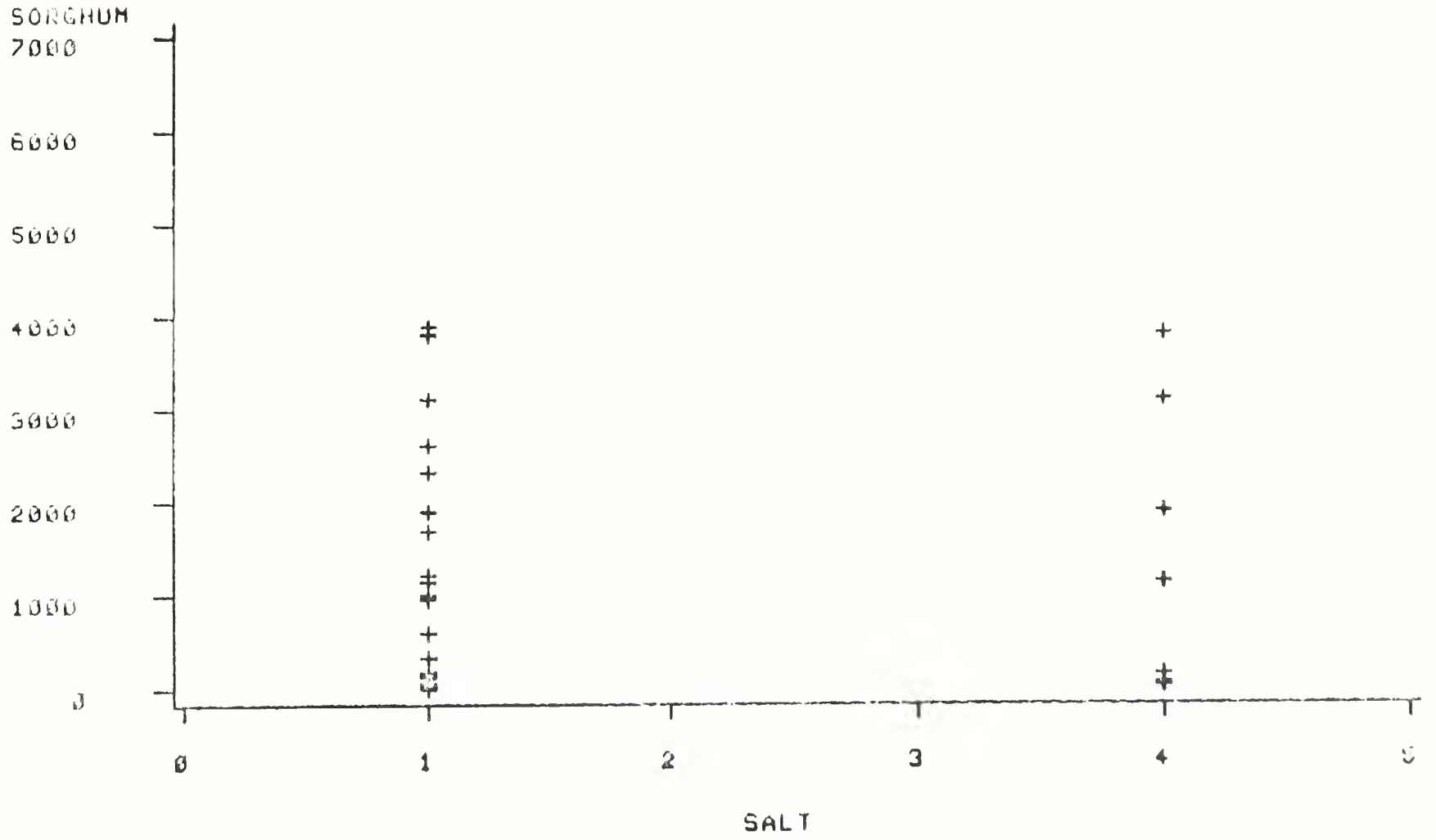


WHEAT VS. DRAINAGE

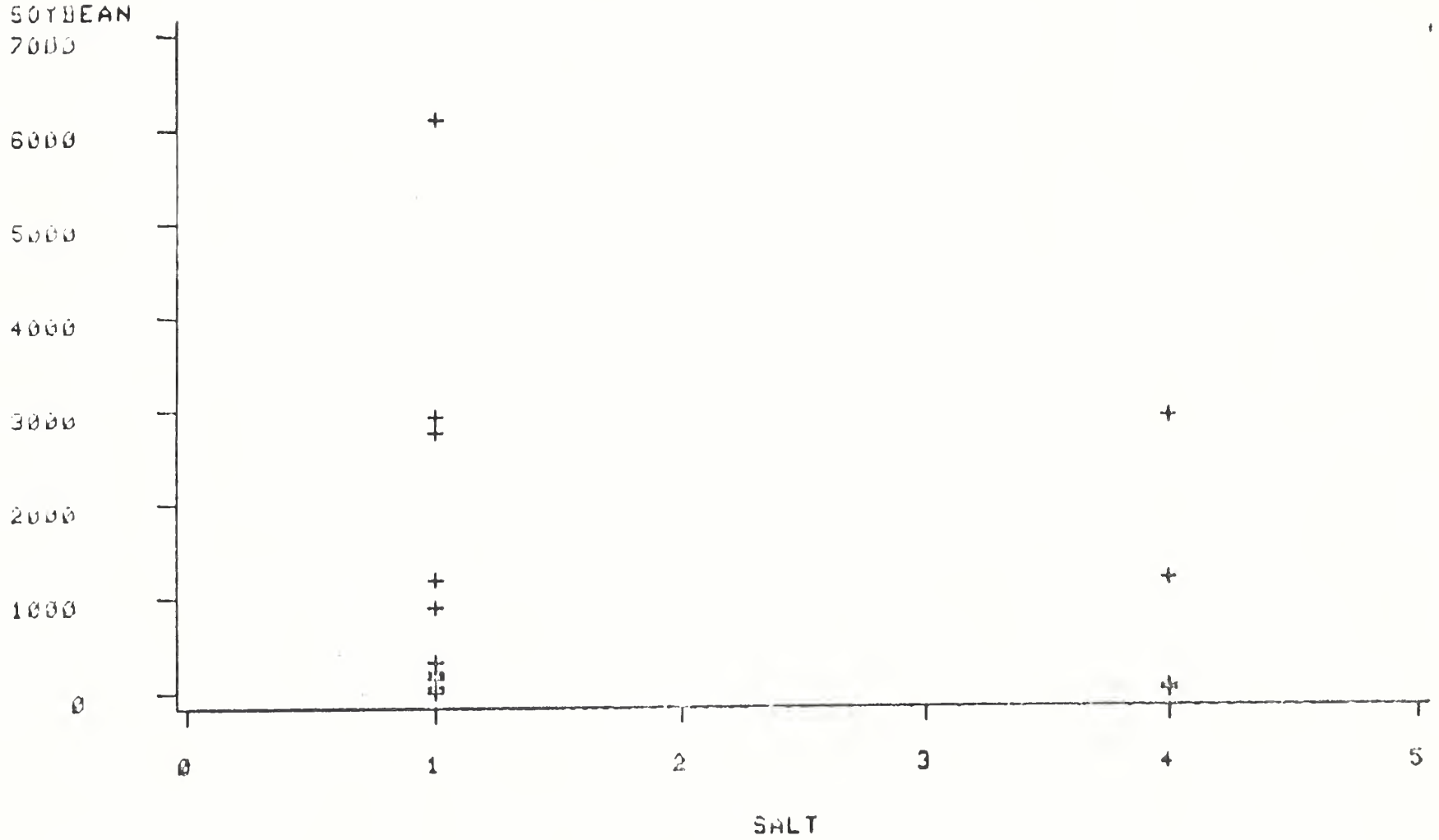


SEP 23 1982

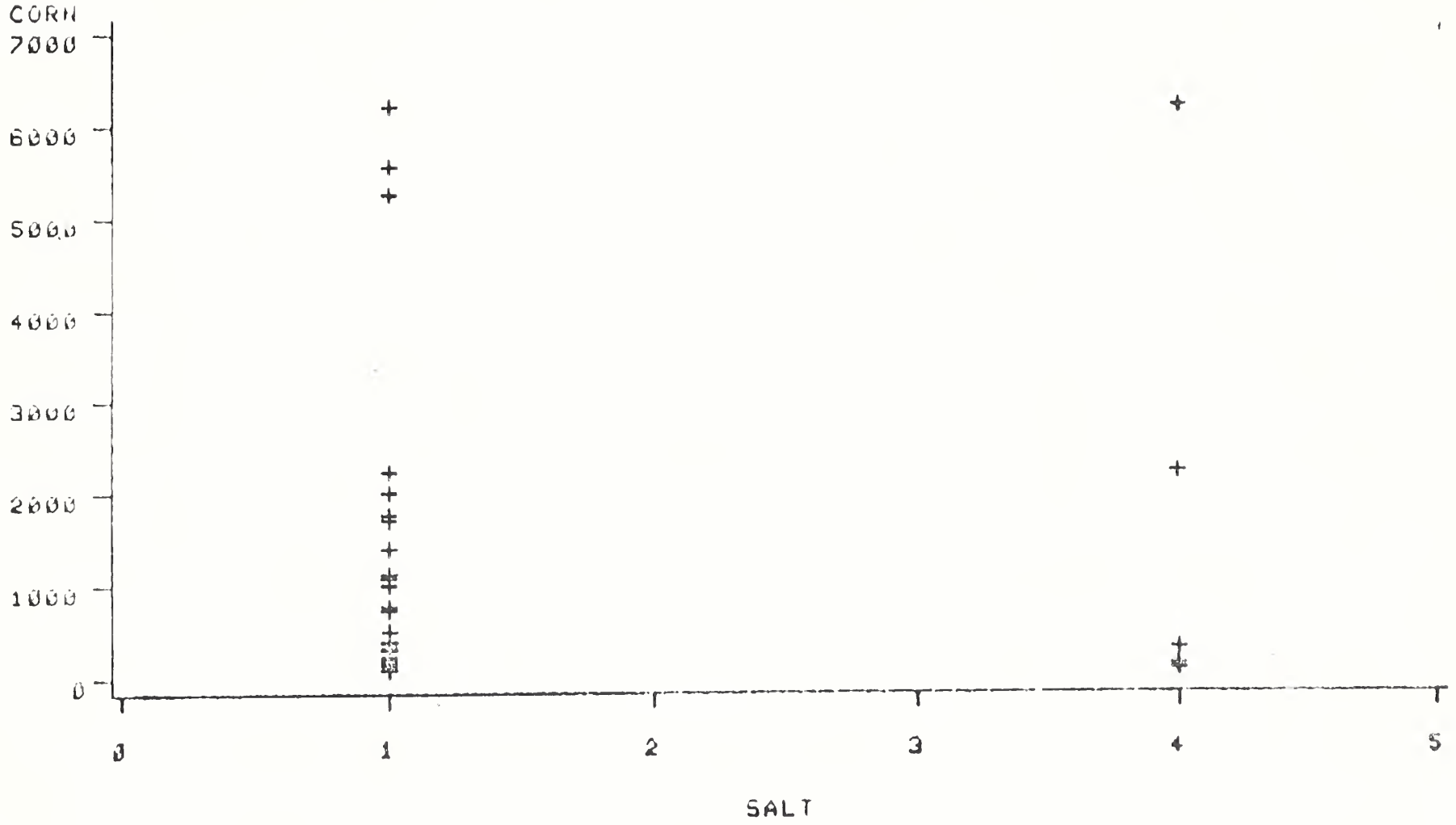
SORGHUM VS. SALT



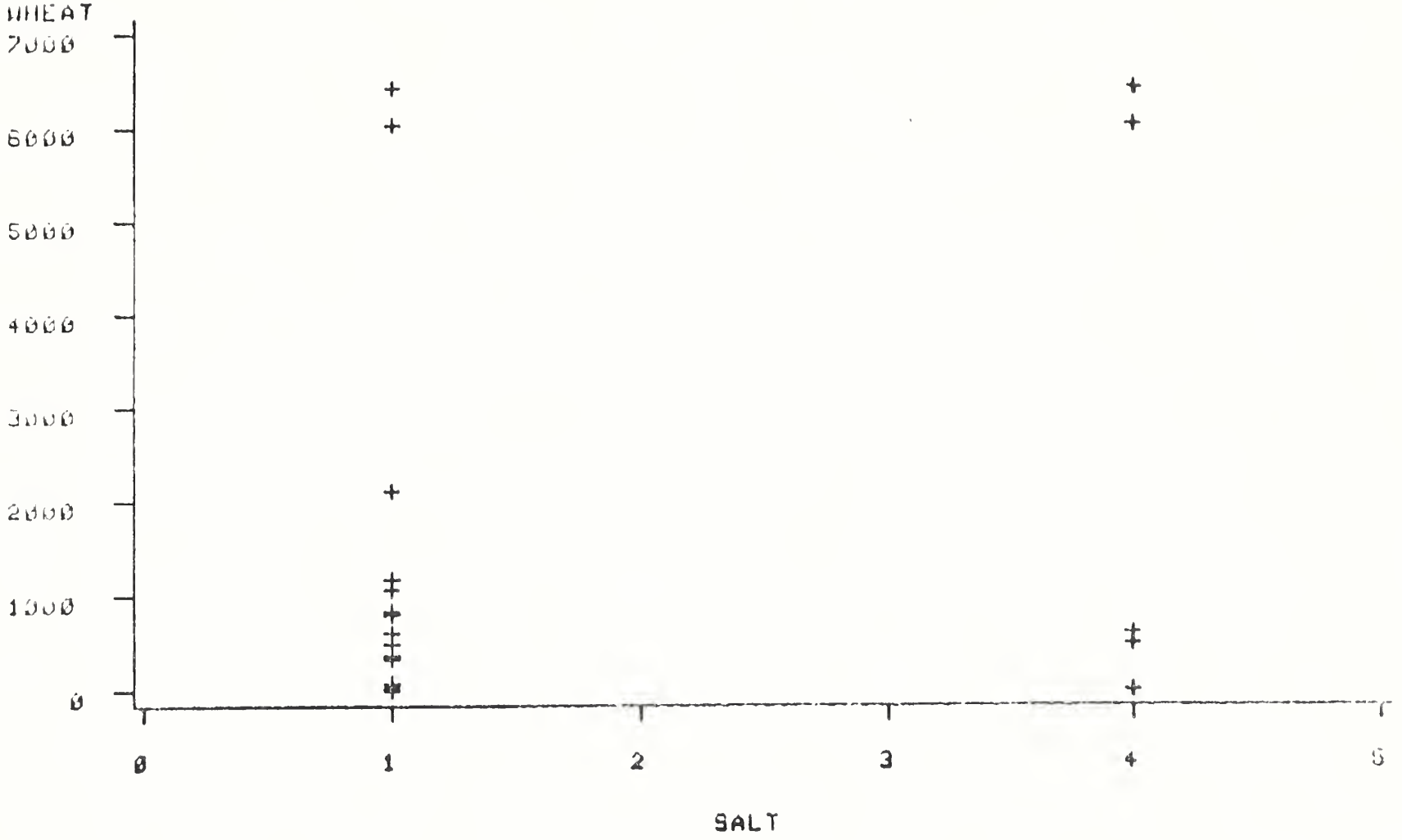
SOYBEAN VS. SALT



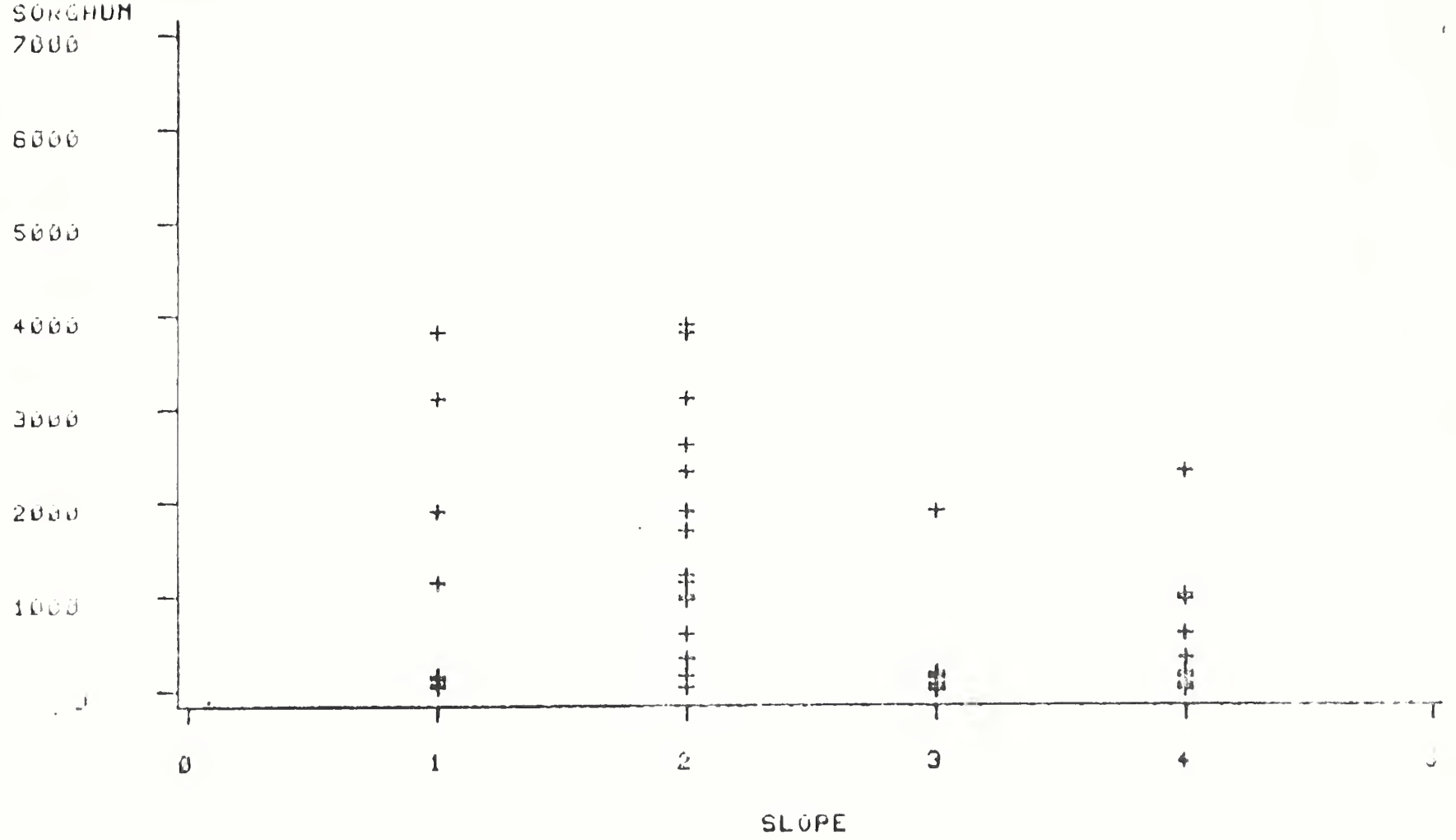
CORN VS. SALT



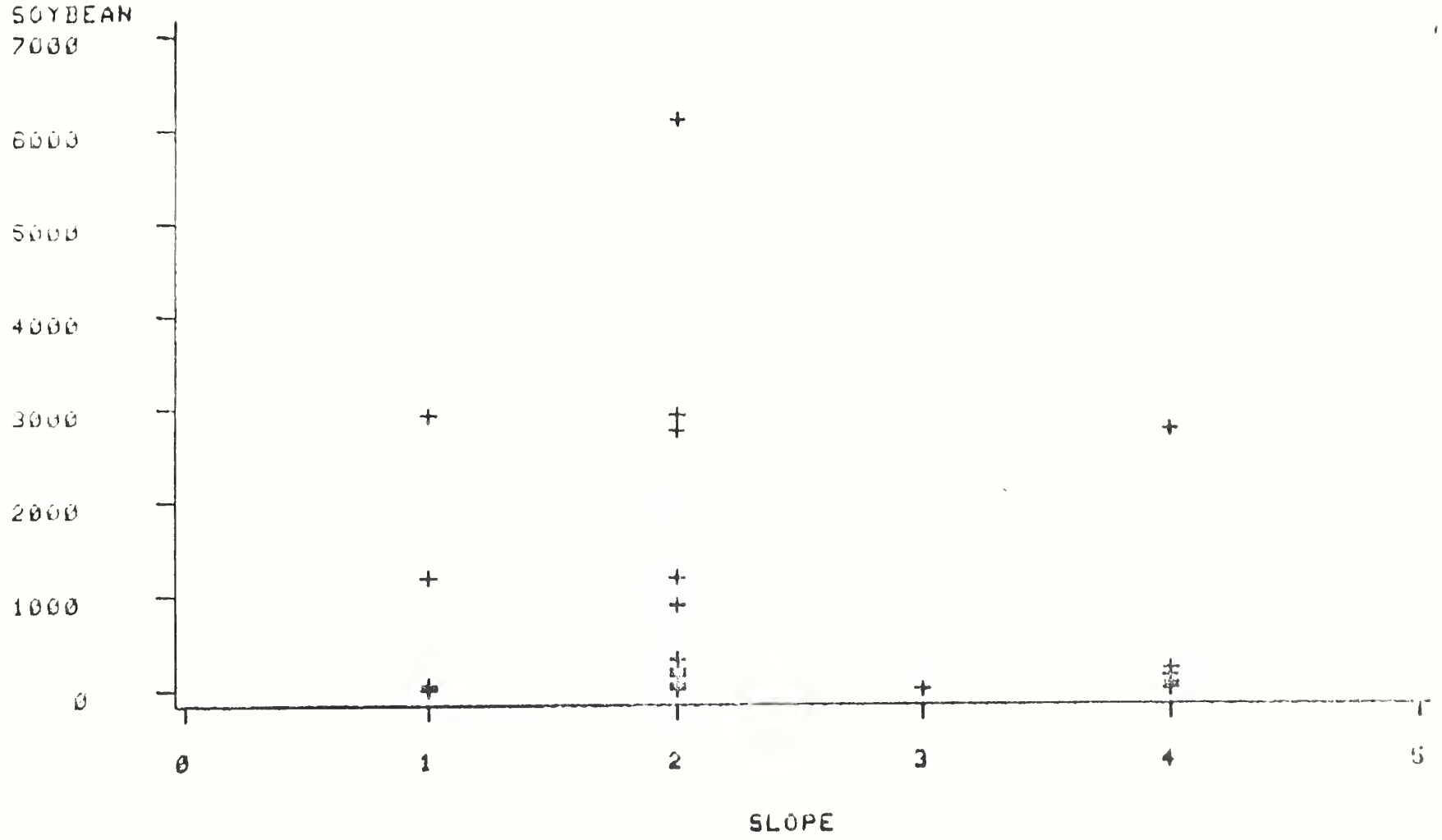
WHEAT VS. SALT



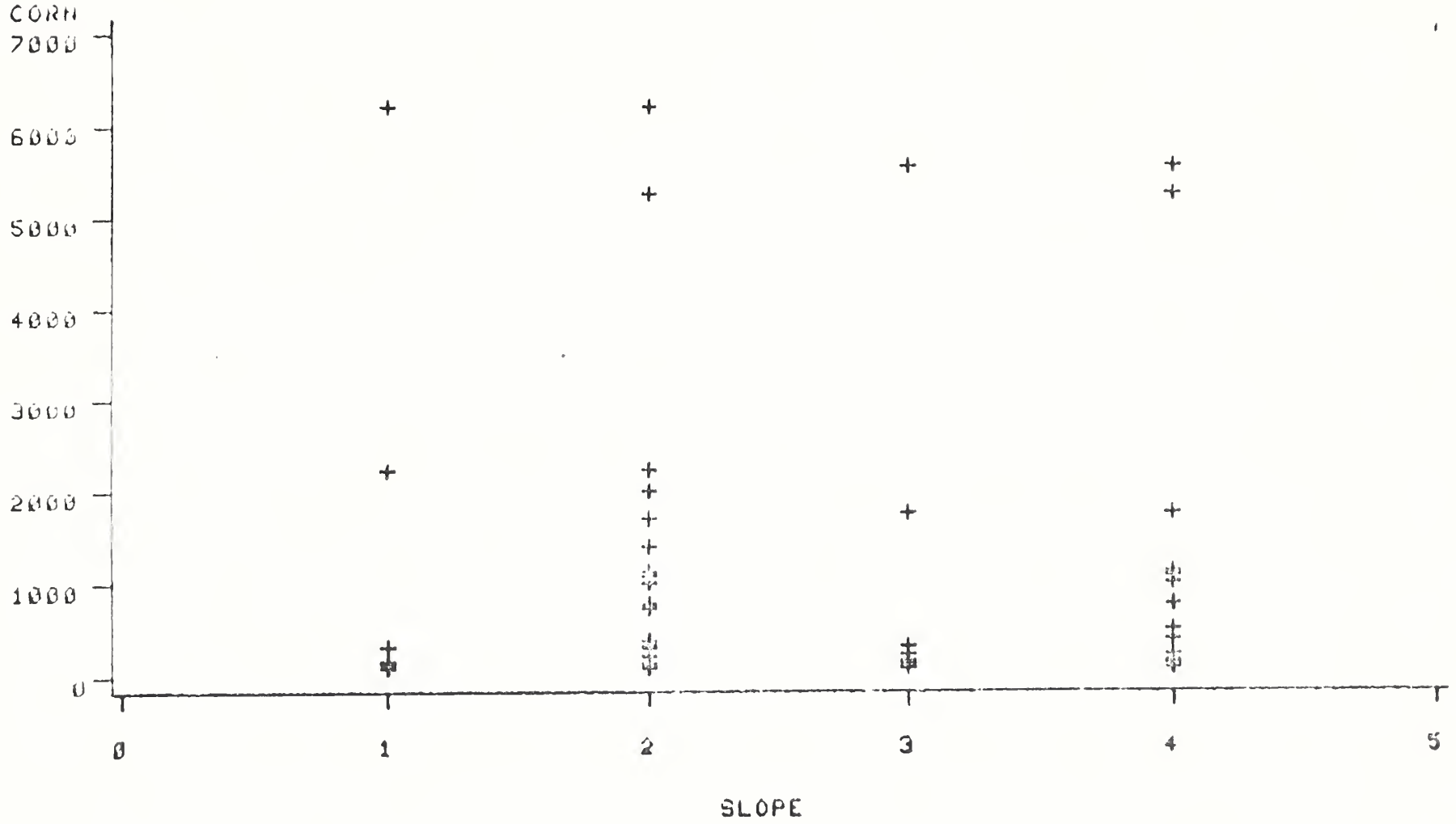
SORGHUM VS. SLOPE

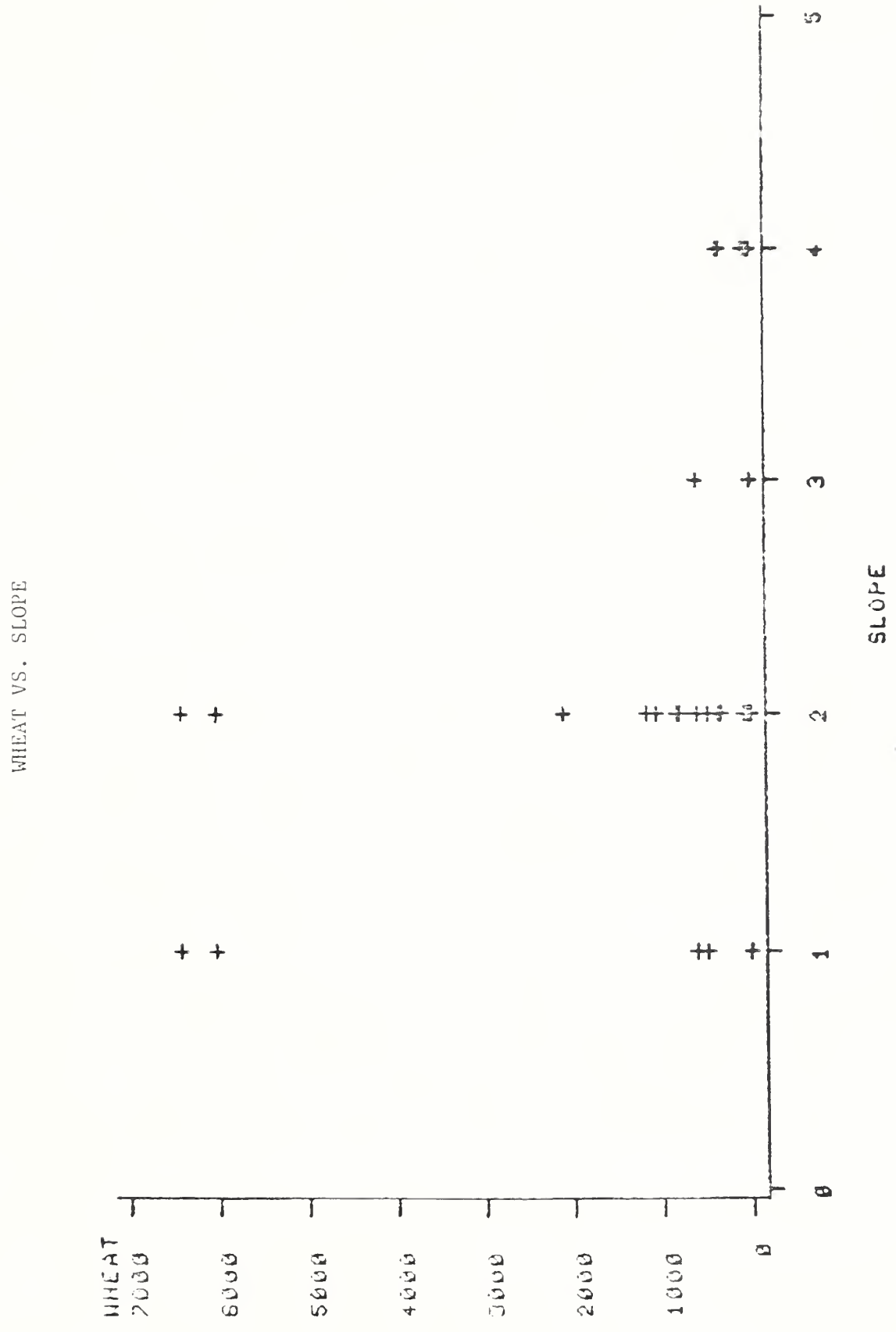


SOYBEAN VS. SLOPE



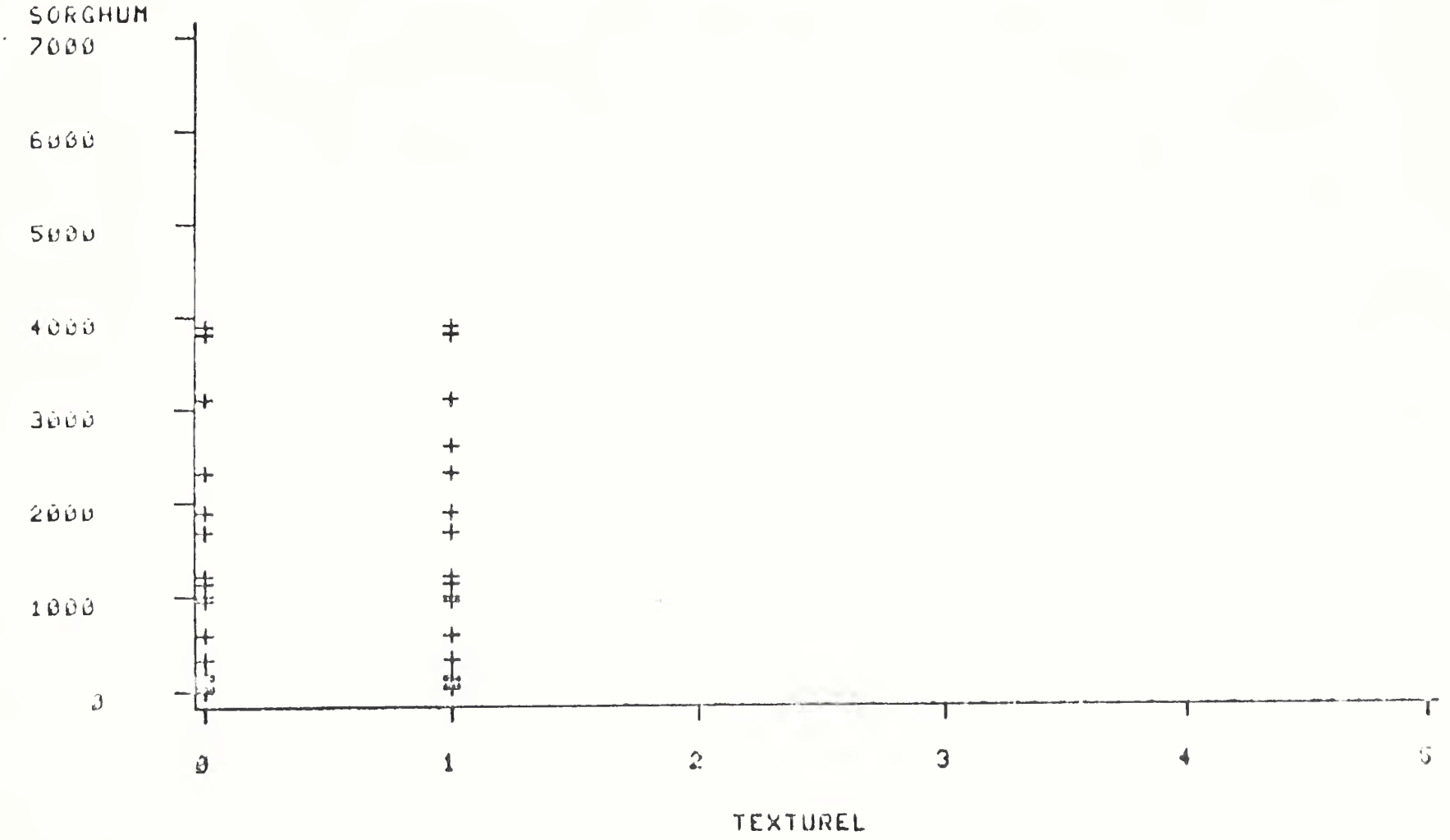
CORN VS. SLOPE




WHEAT VS. SLOPE
WHEAT
7000
6000
5000
4000
3000
2000
1000
0
0
1
2
3
5
SLOPE

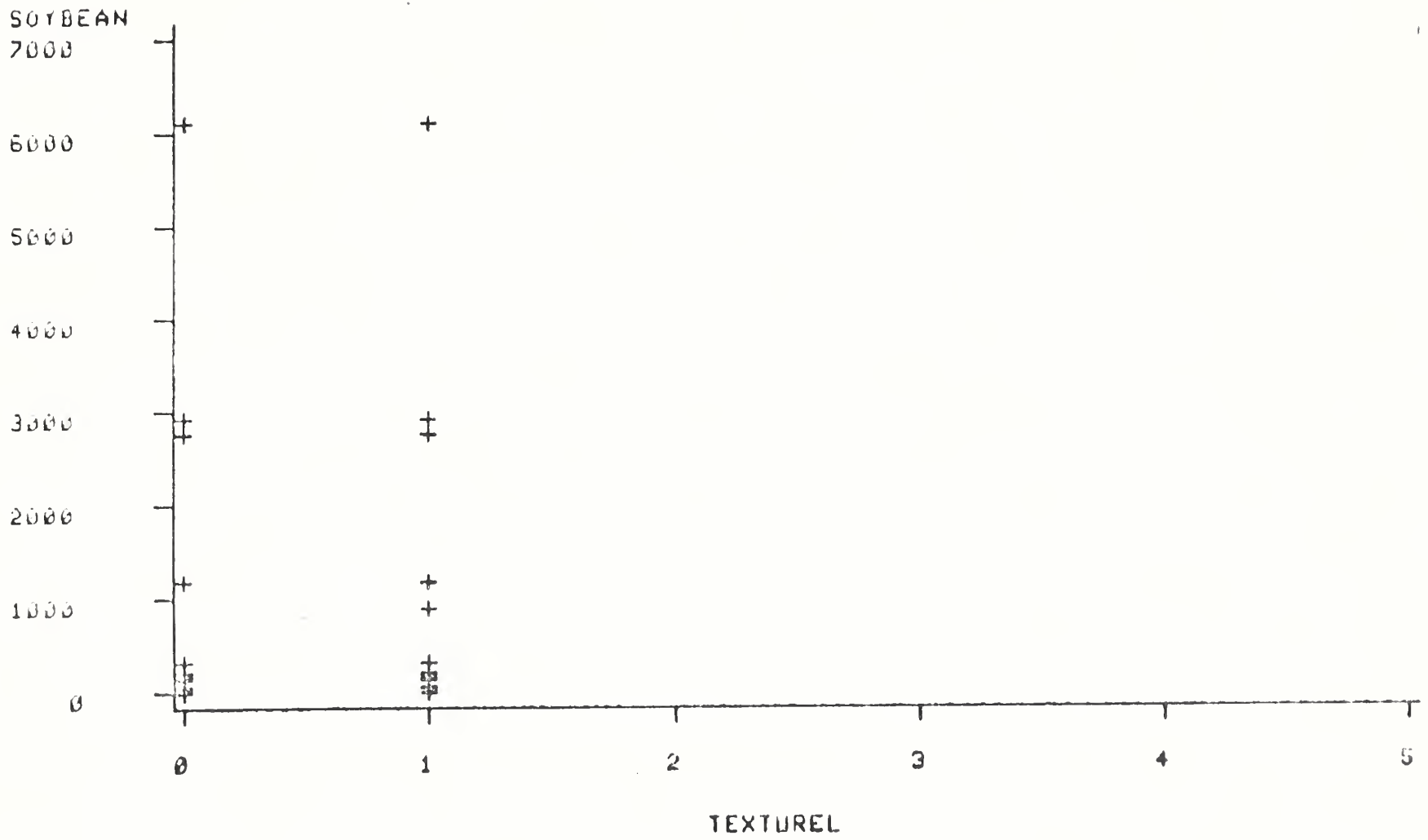
SEP 23 1982

SORGHUM VS. DUMMY TEXTURE

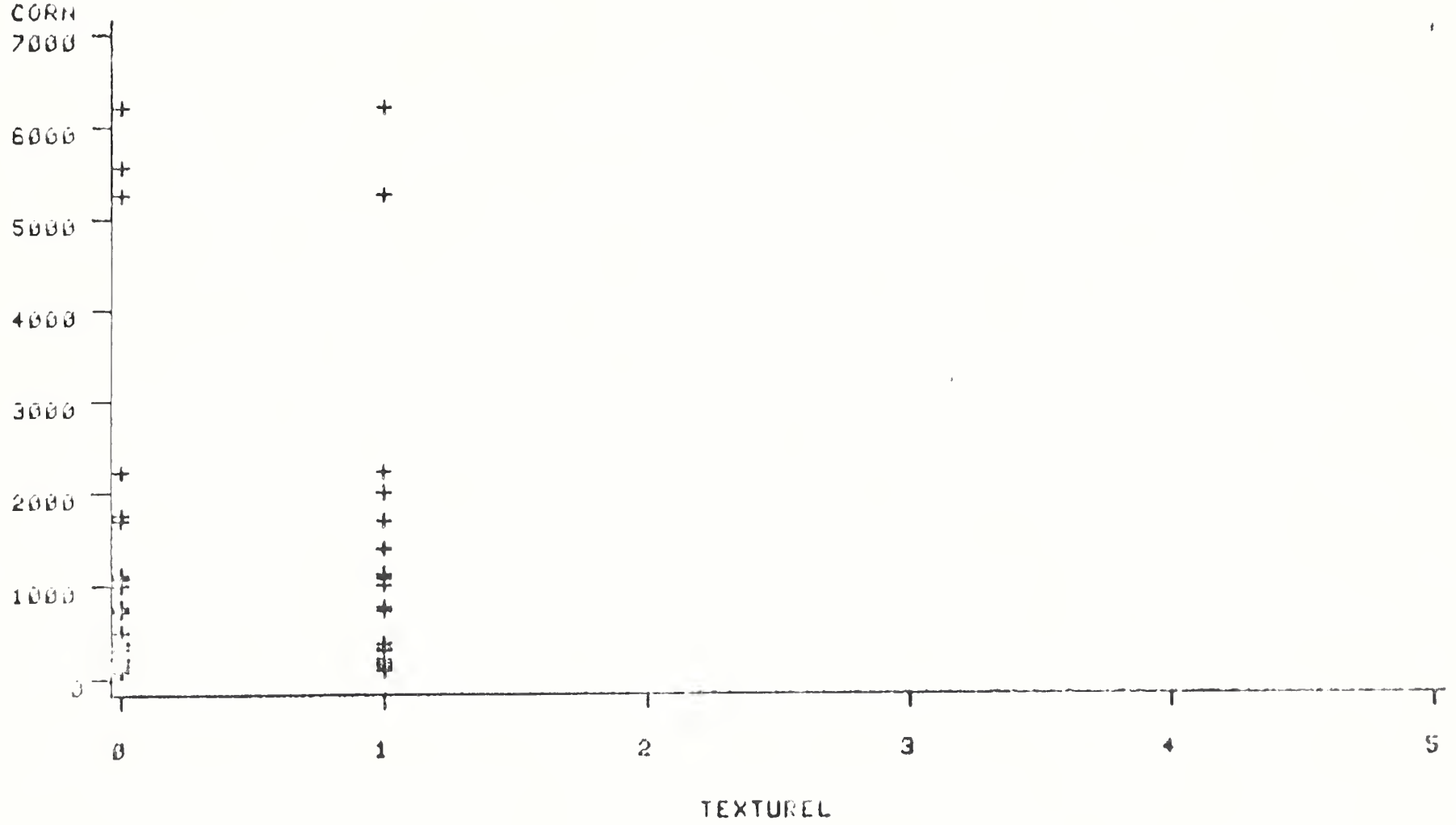


SEP 2 3 1982

SOYBEAN VS. DUMMY TEXTURE

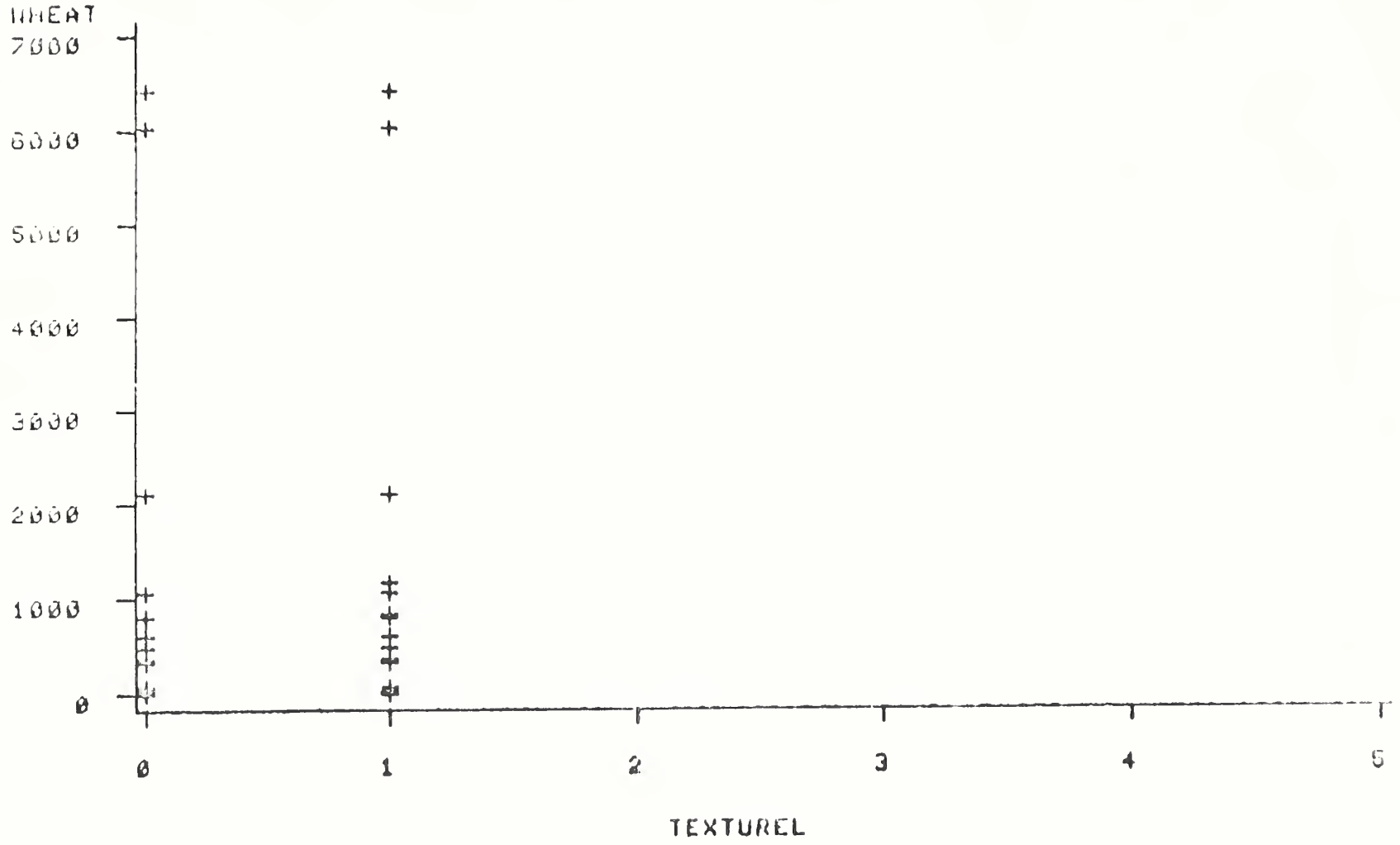


CORN VS. DUMMY TEXTURE

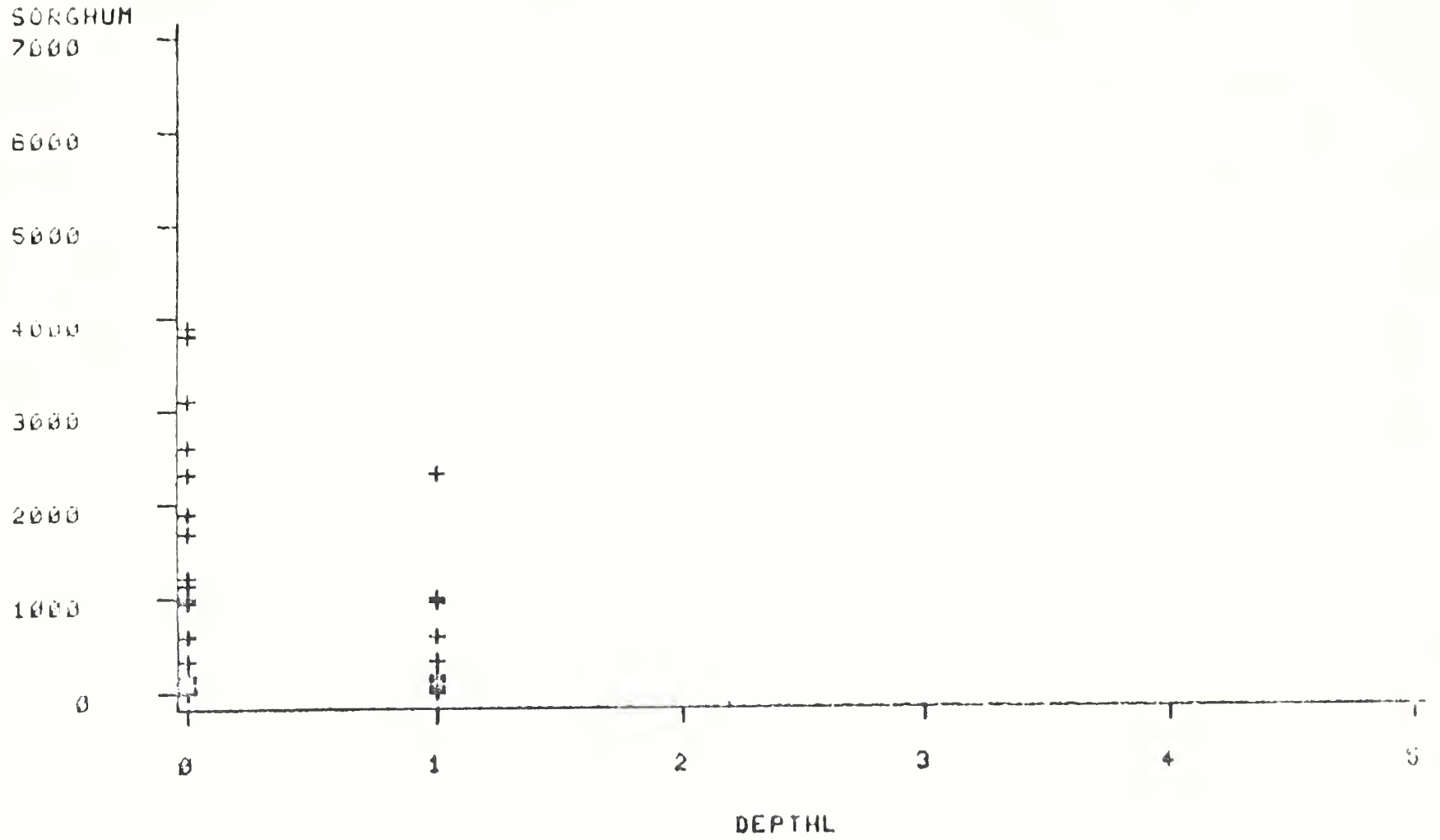


SEP 2 3 1982

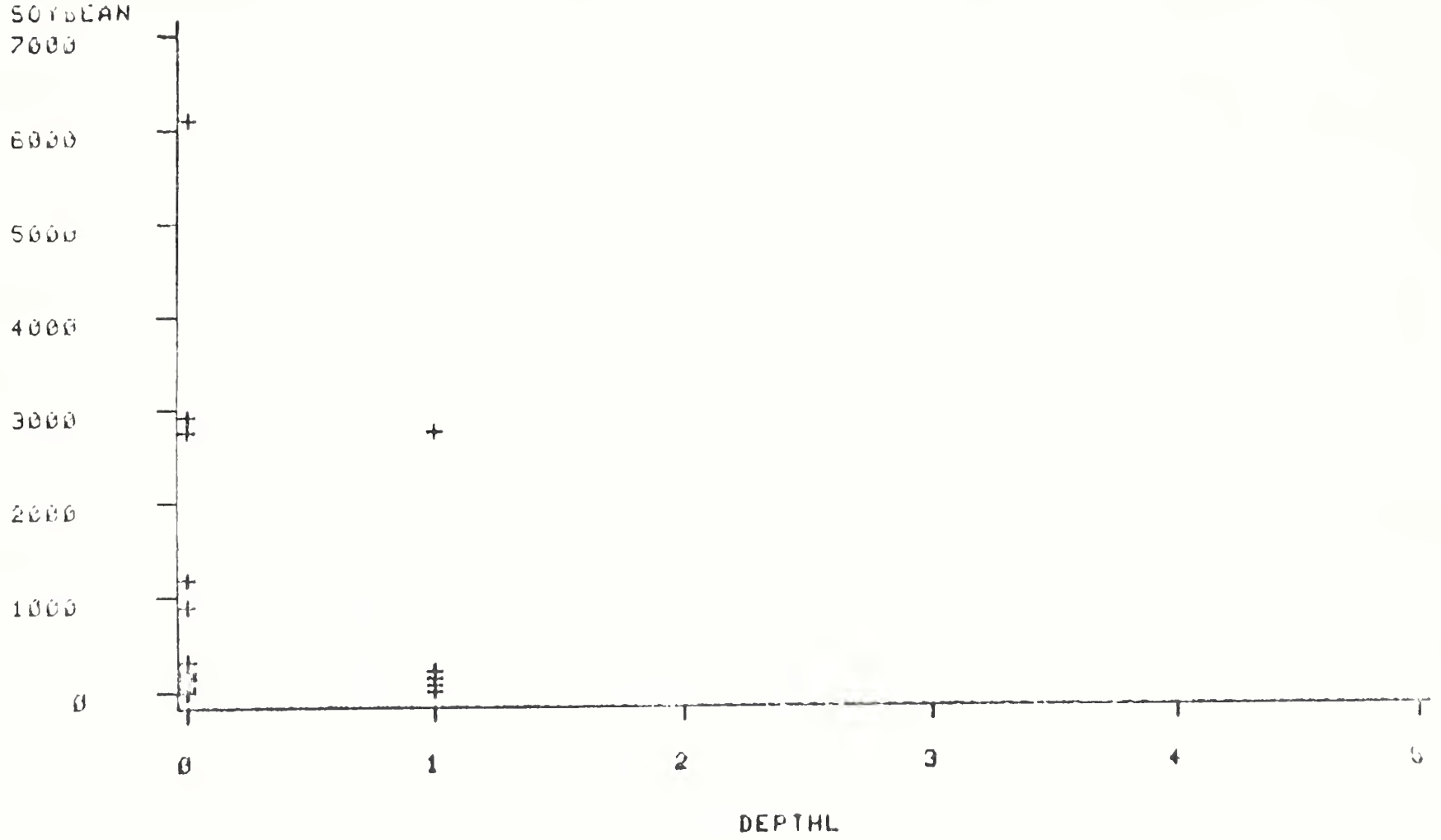
WHEAT VS. DUMMY TEXTURE



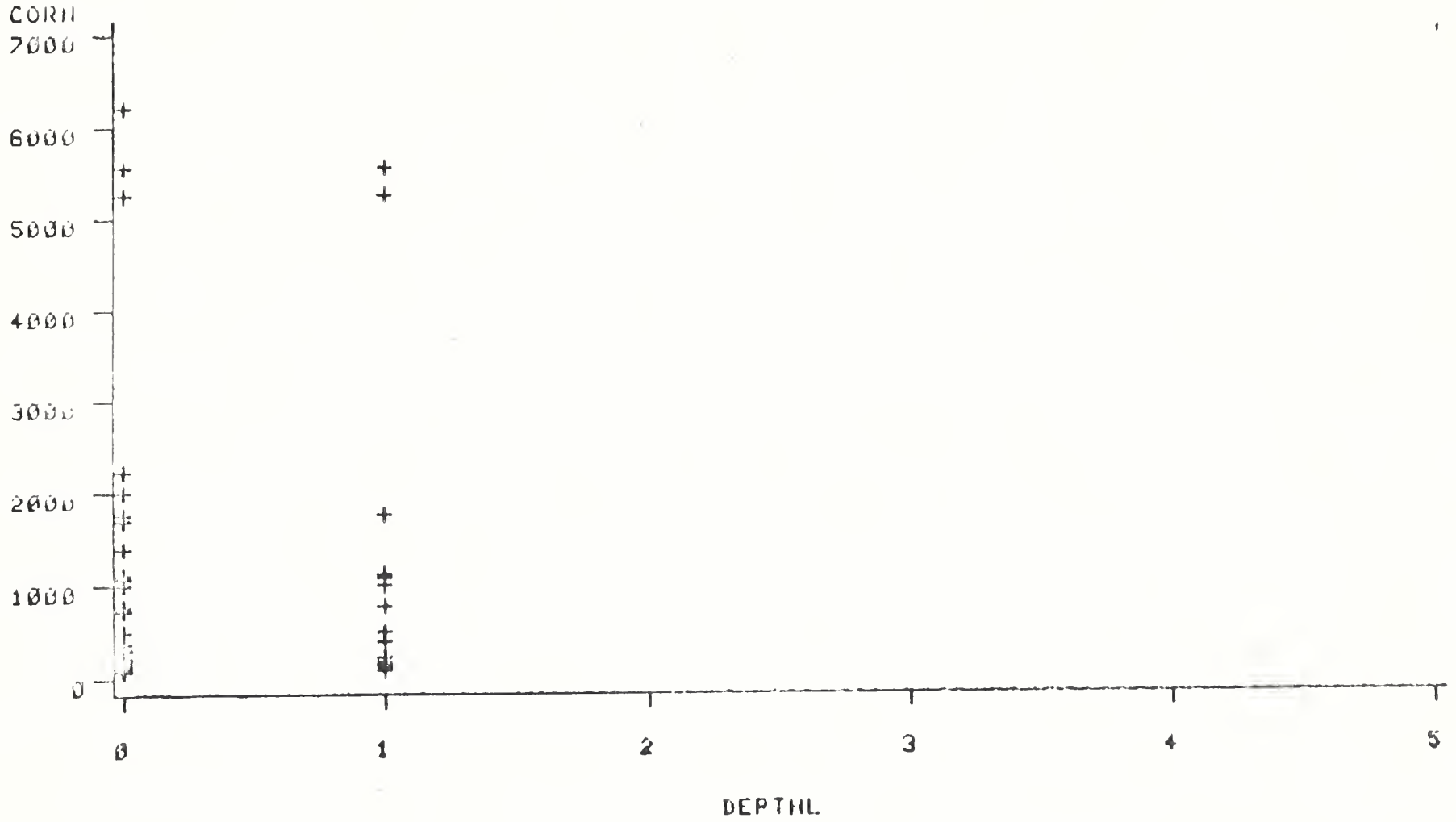
SORGHUM VS. DUMMY DEPTH



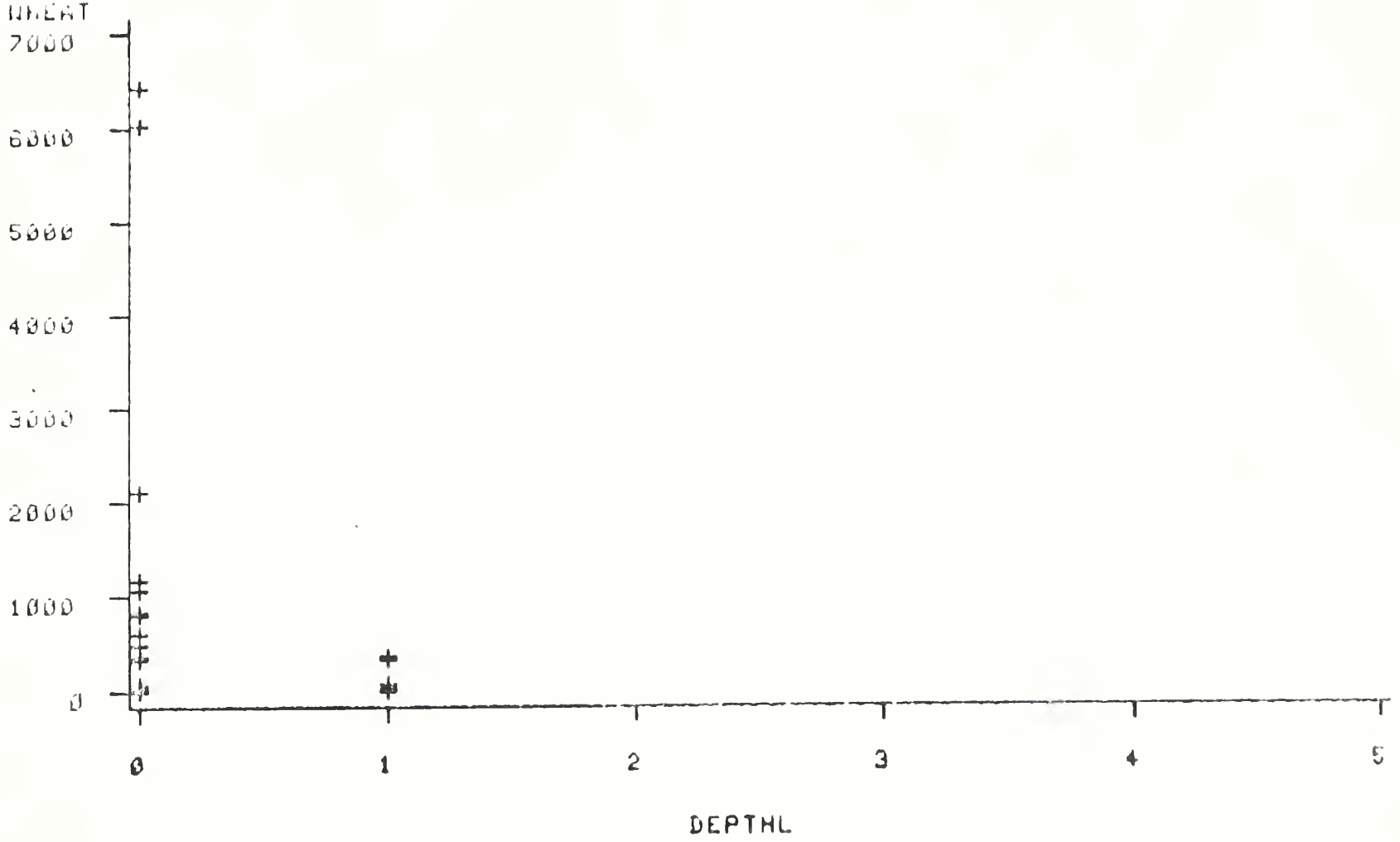
SOYBEAN VS. DUMMY DEPTH



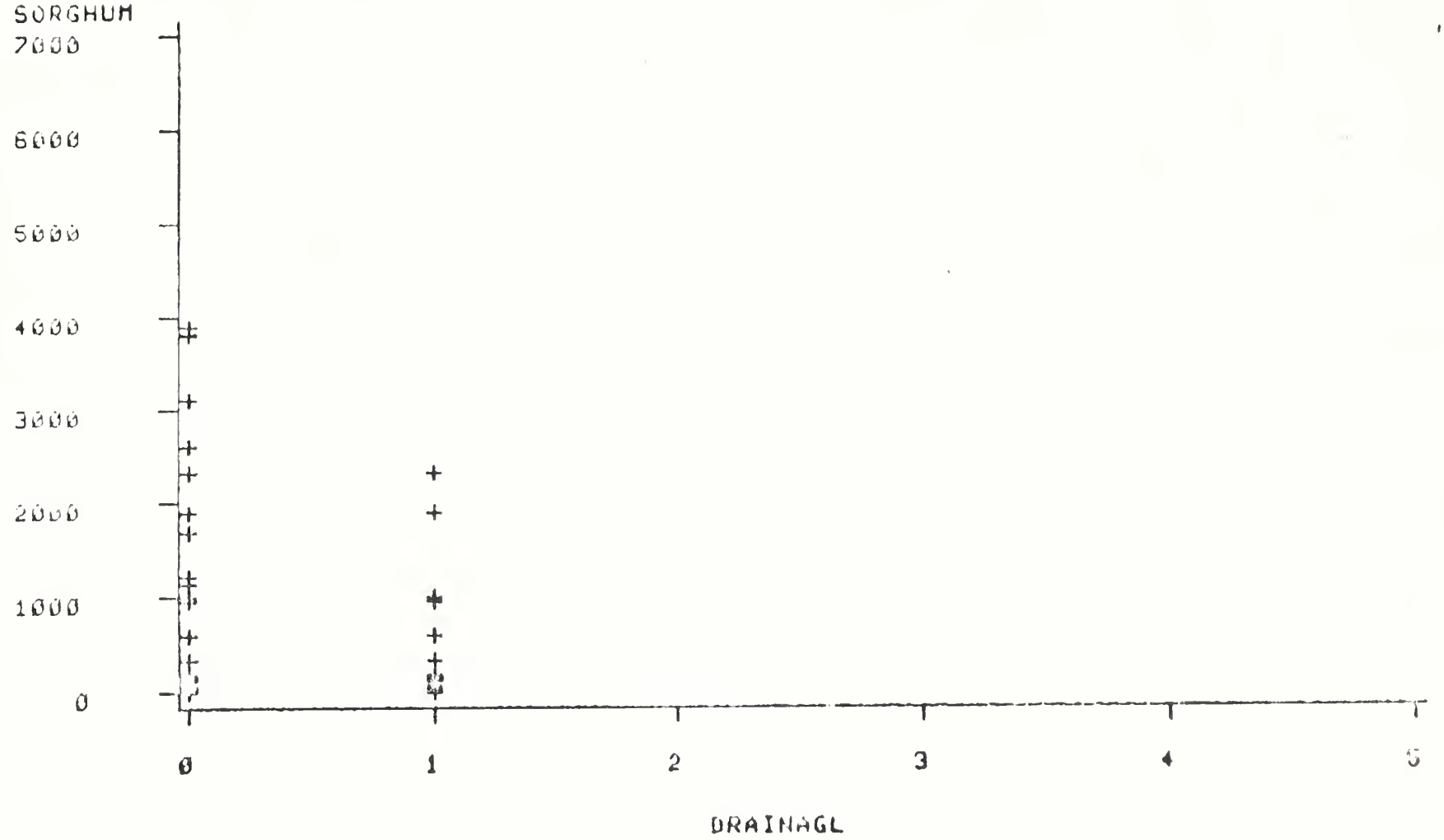
CORN VS. DUMMY DEPTH



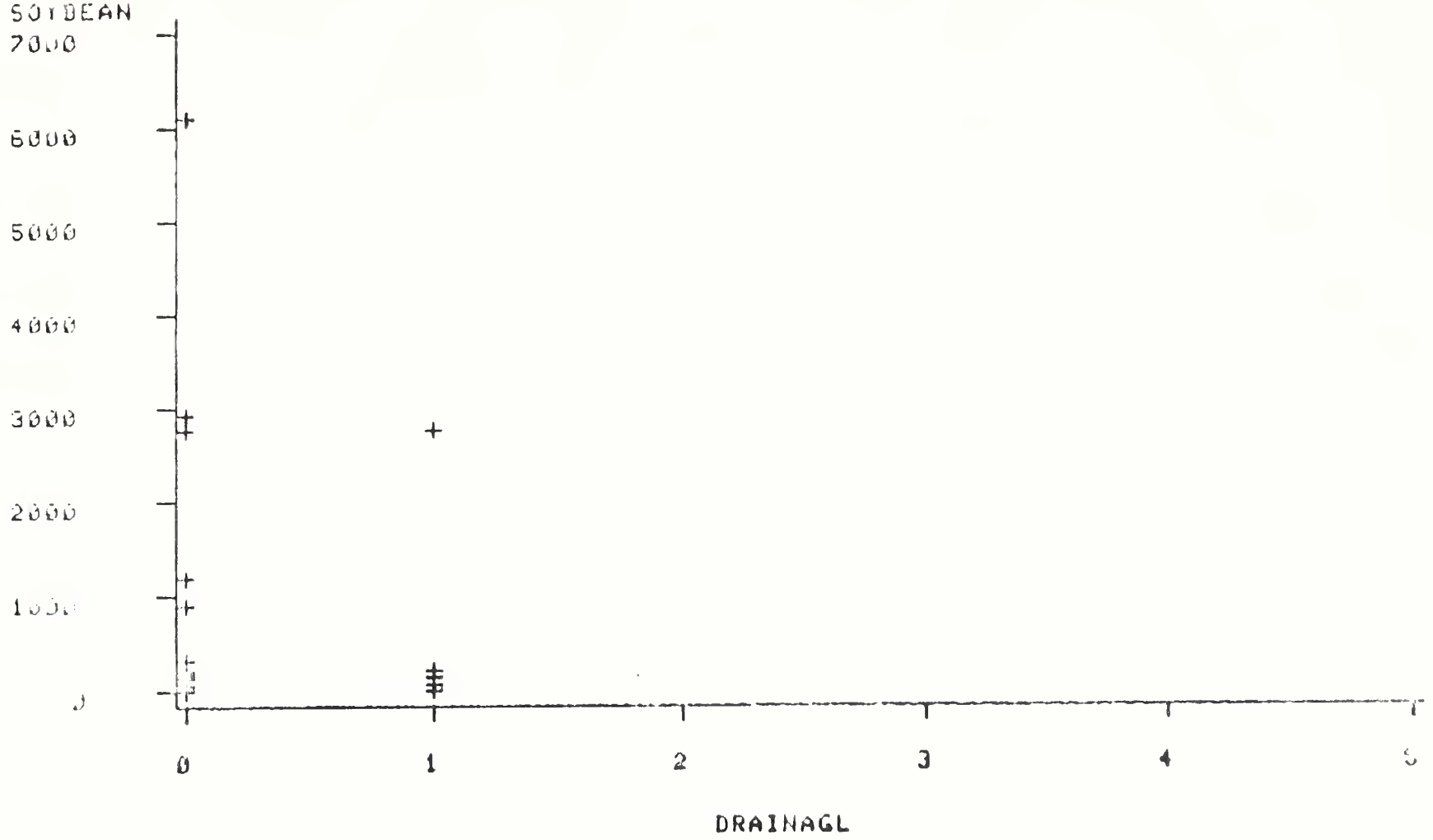
WHEAT VS. DUMMY DEPTH



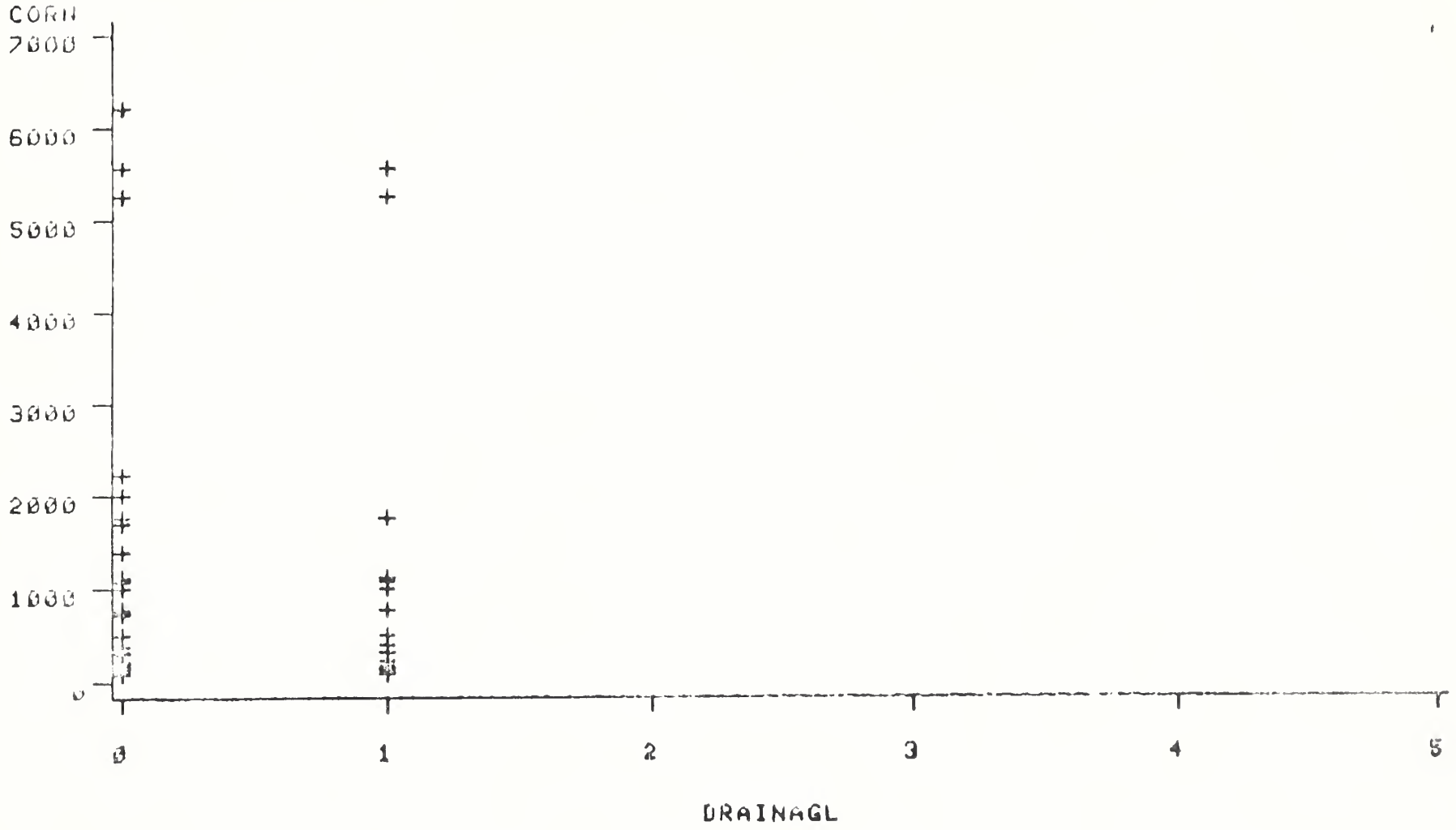
SORGHUM VS. DUMMY DRAINAGE



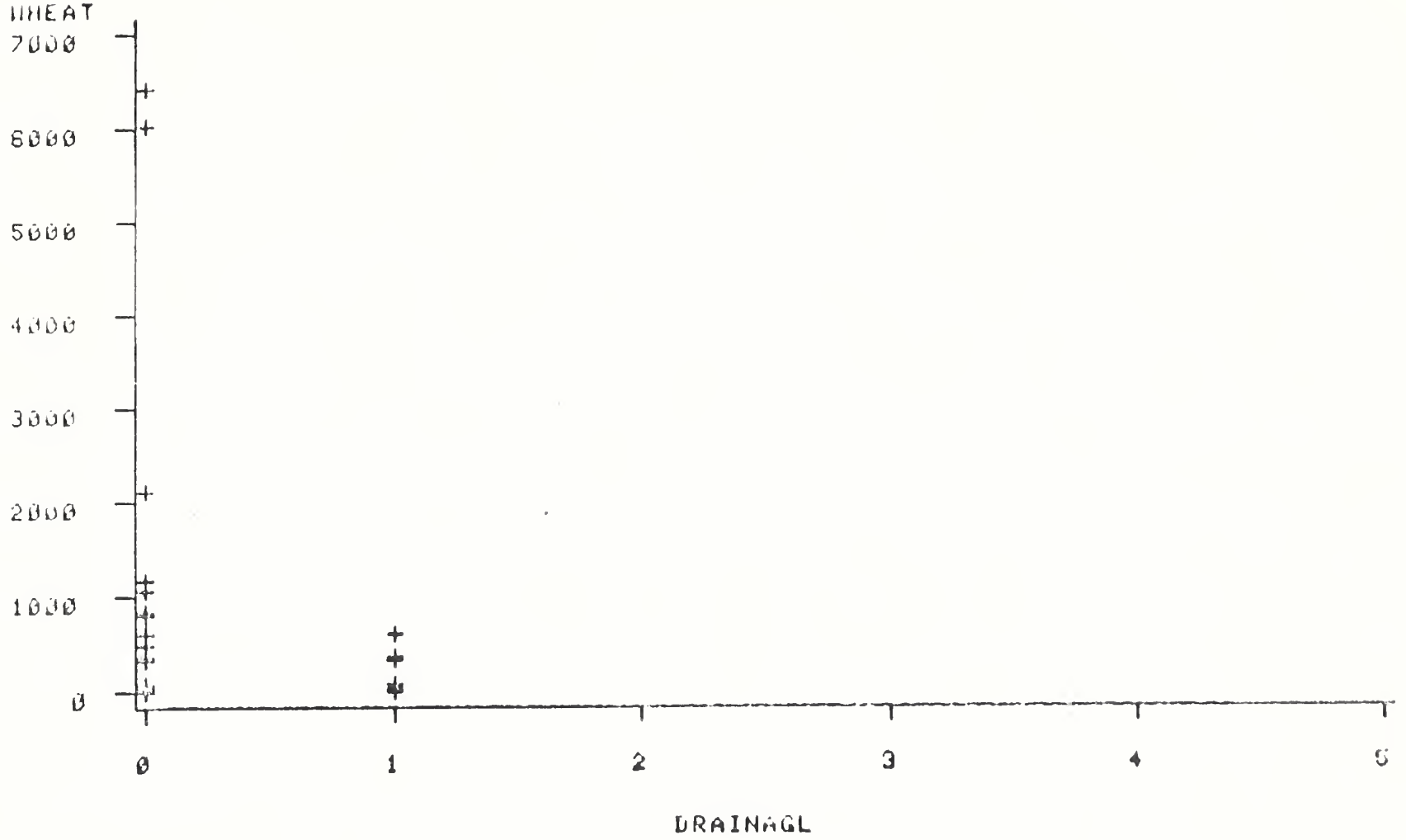
SOYBEAN VS. DUMMY DRAINAGE



CORN VS. DUMMY DRAINAGE

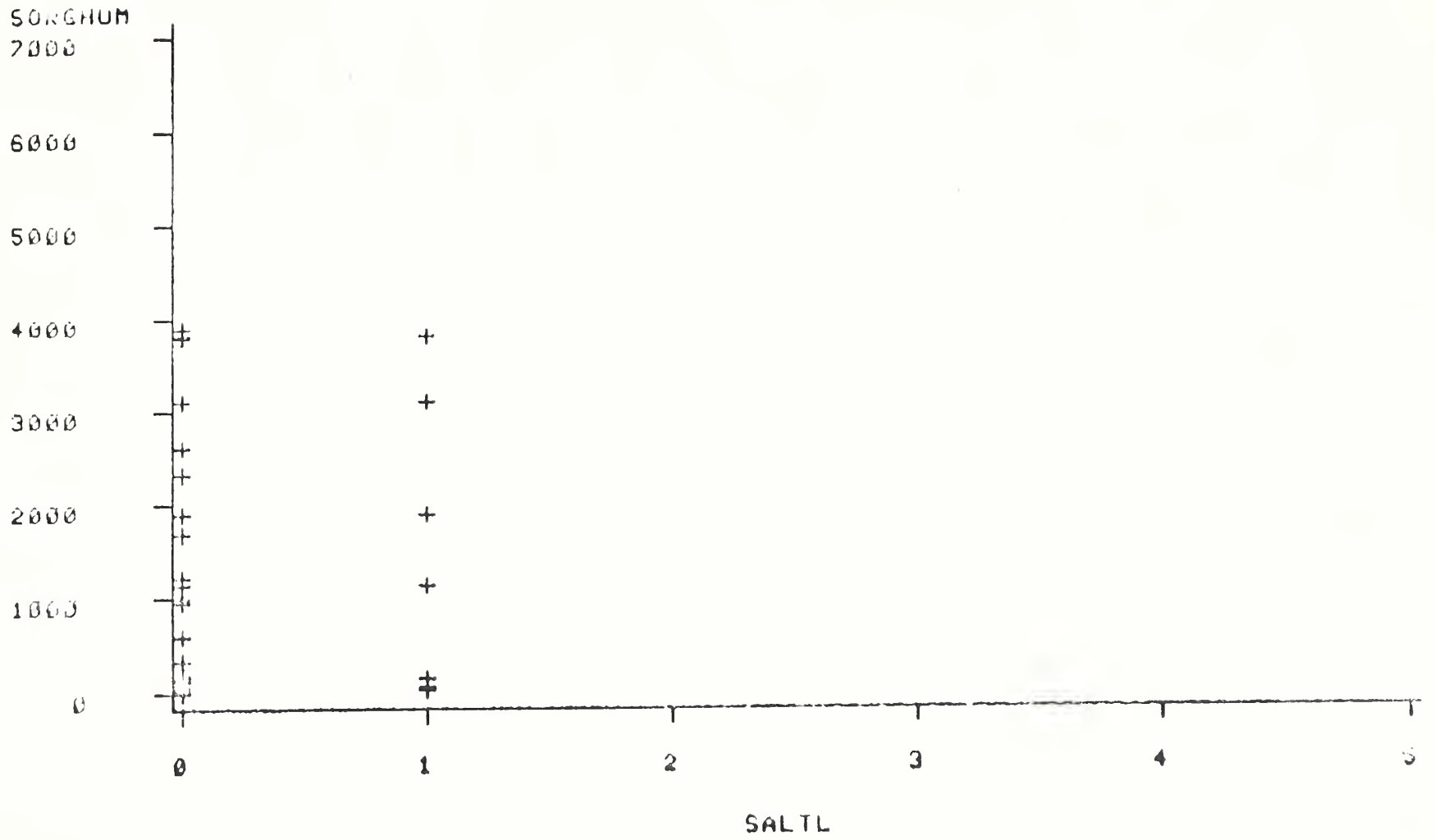


WHEAT VS. DUMMY DRAINAGE

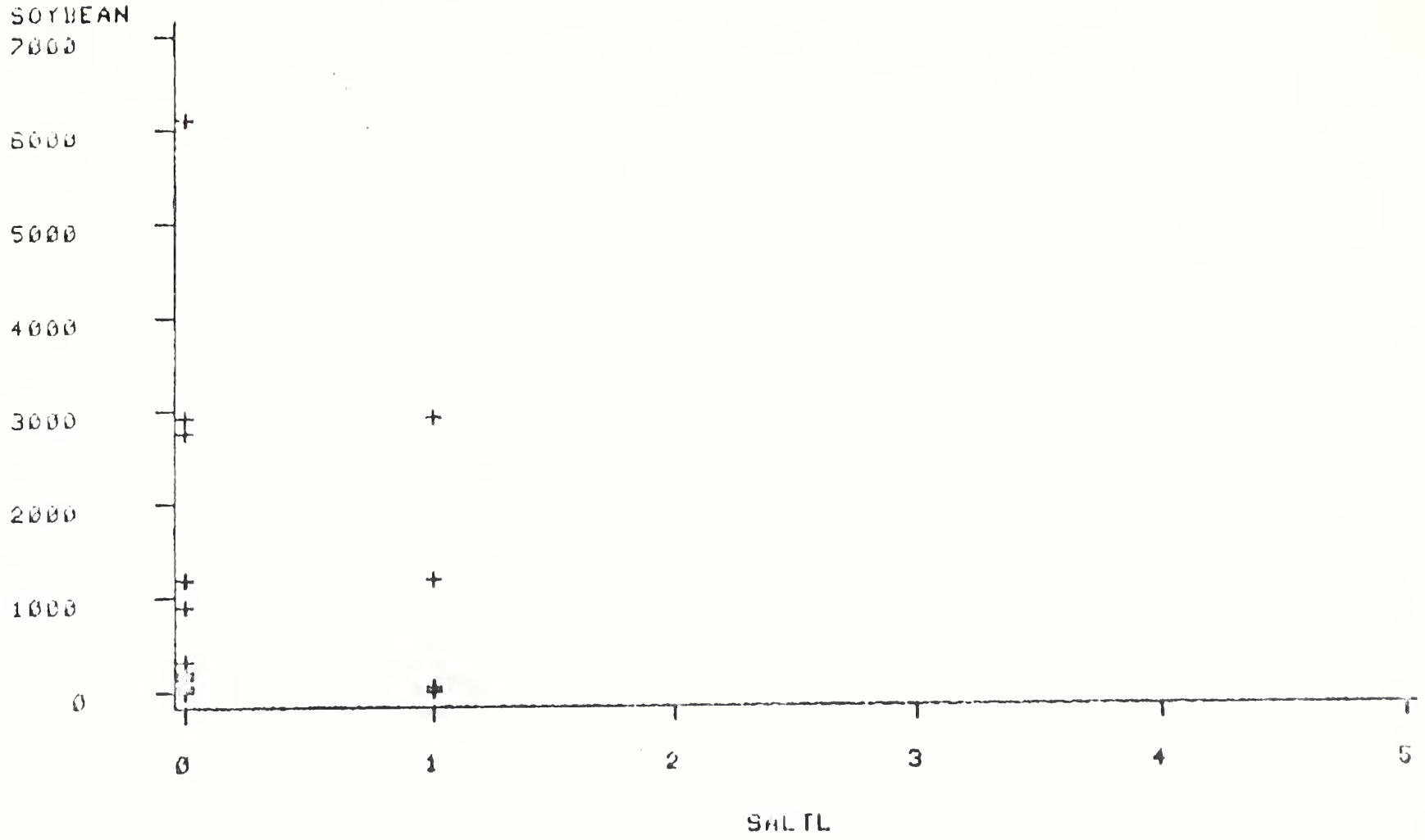


SEP 23 1982

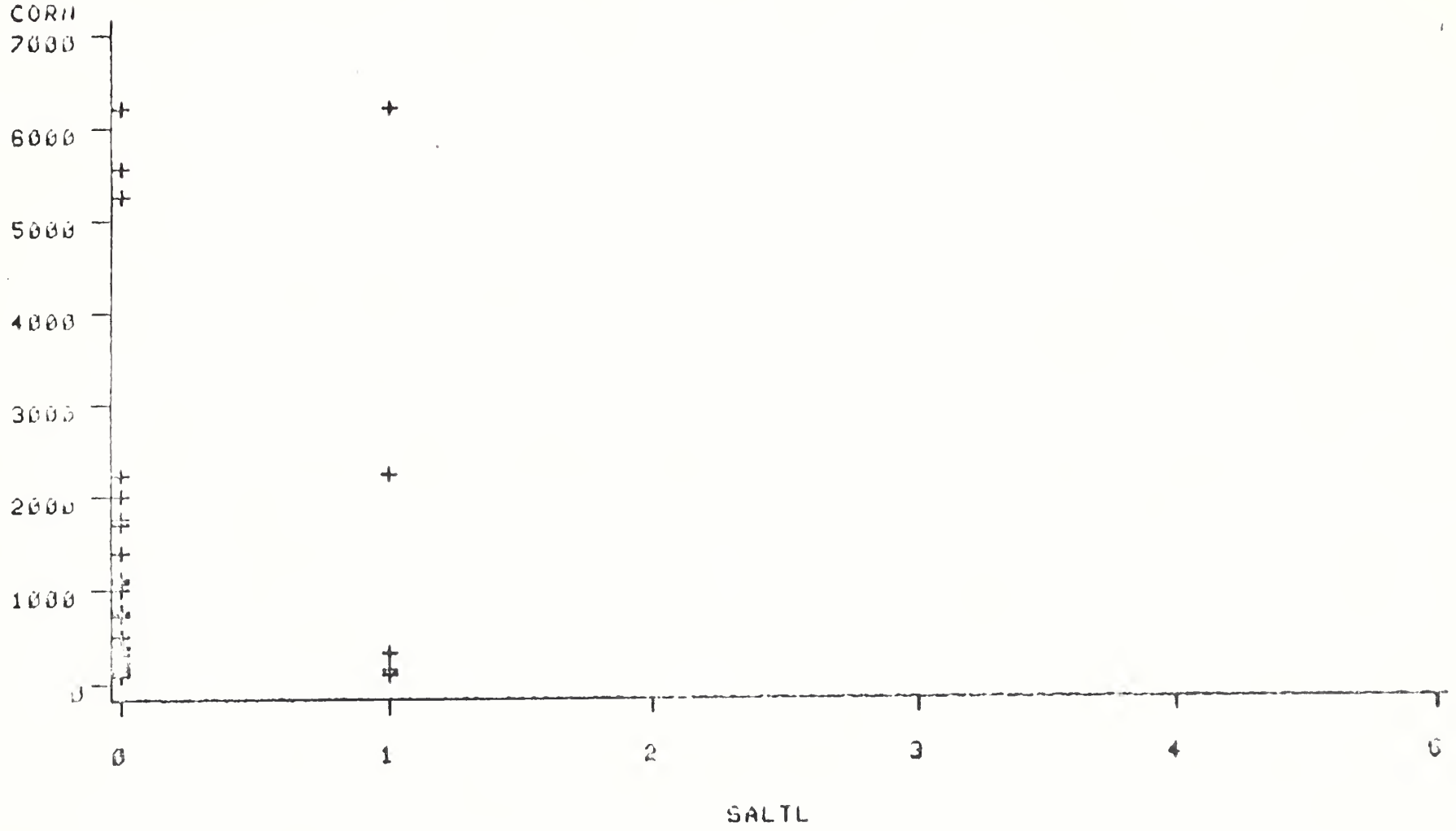
SORGHUM VS. DUMMY SALT



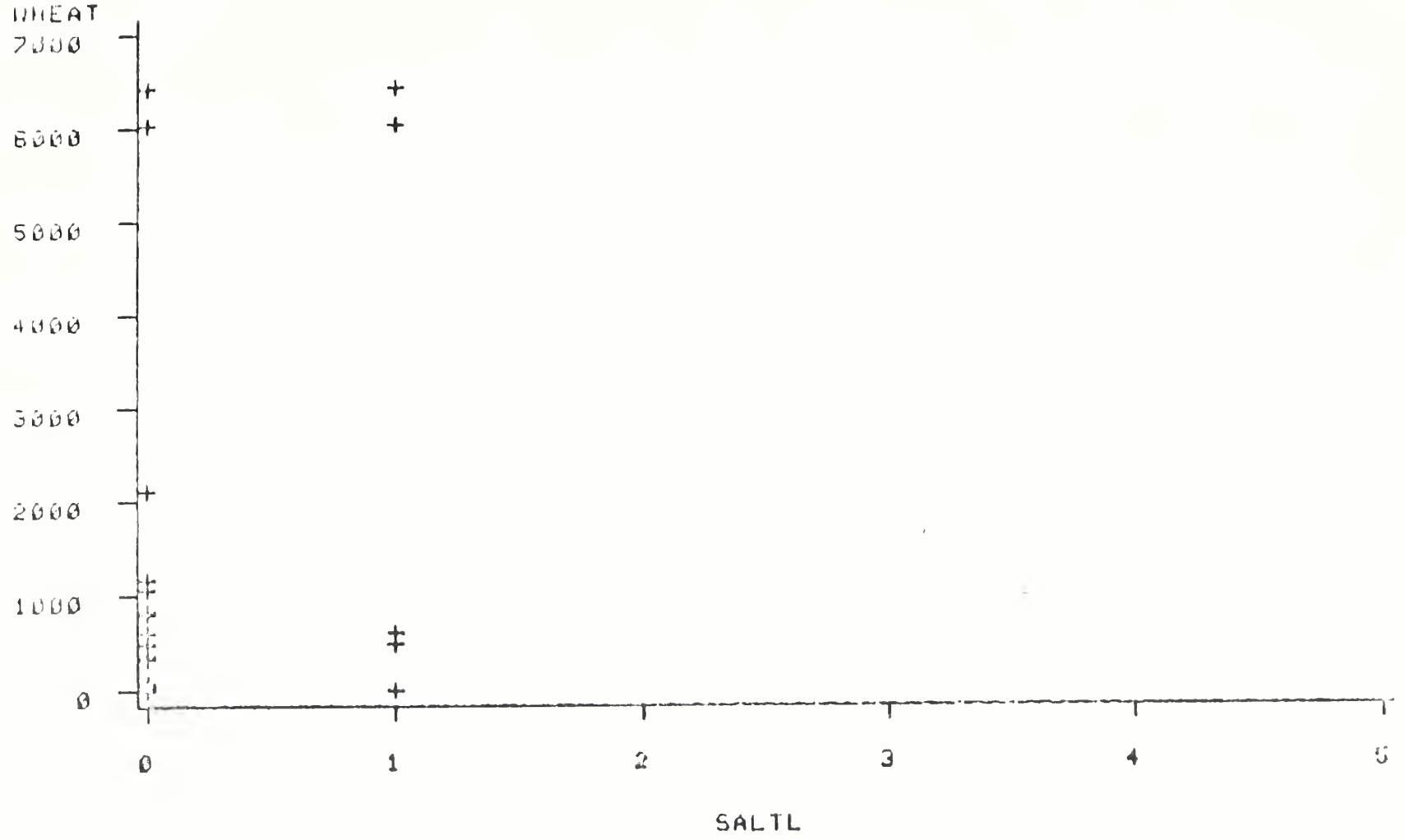
SOYBEAN VS. DUMMY SALT



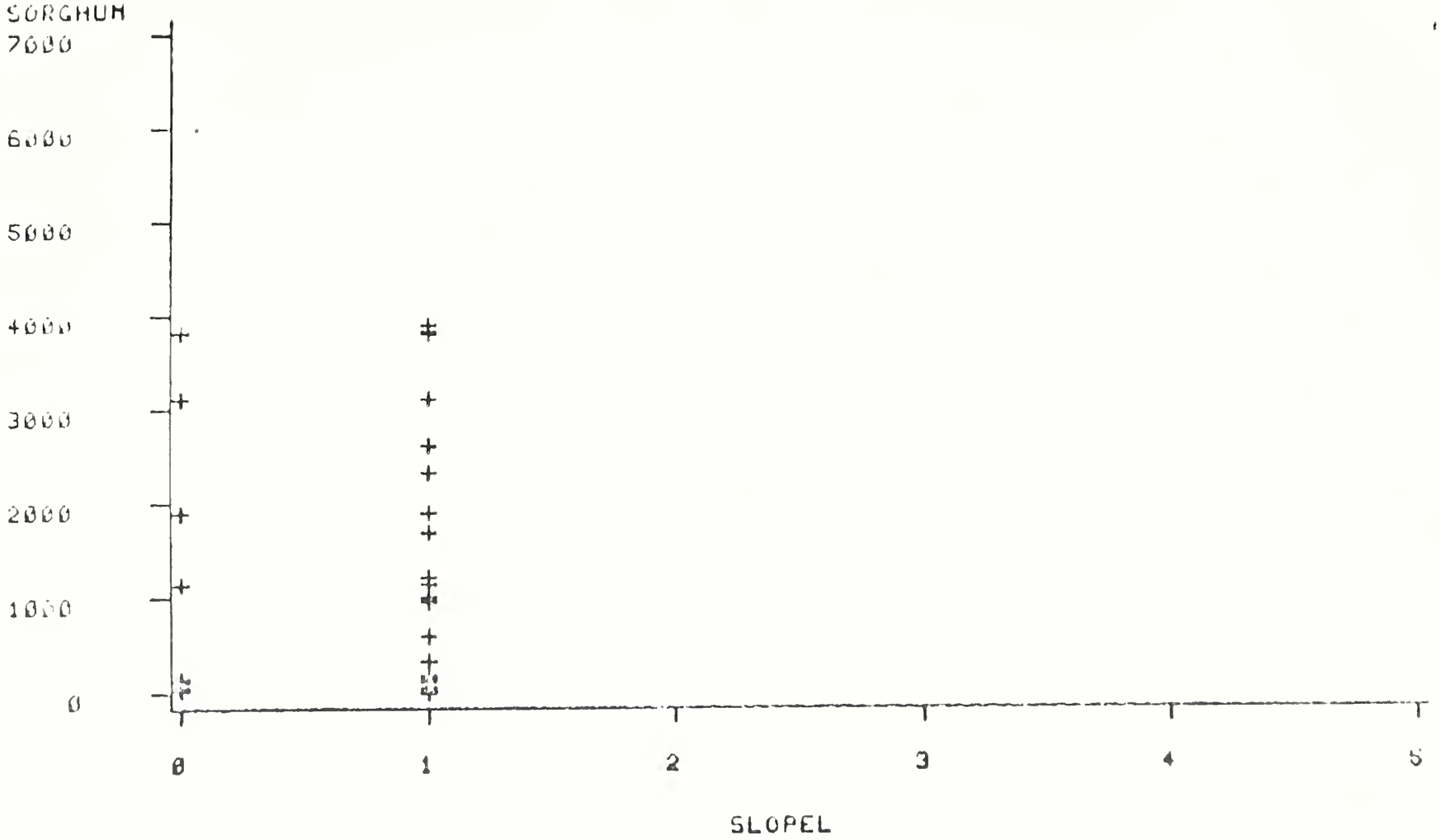
CORN VS. DUMMY SALT



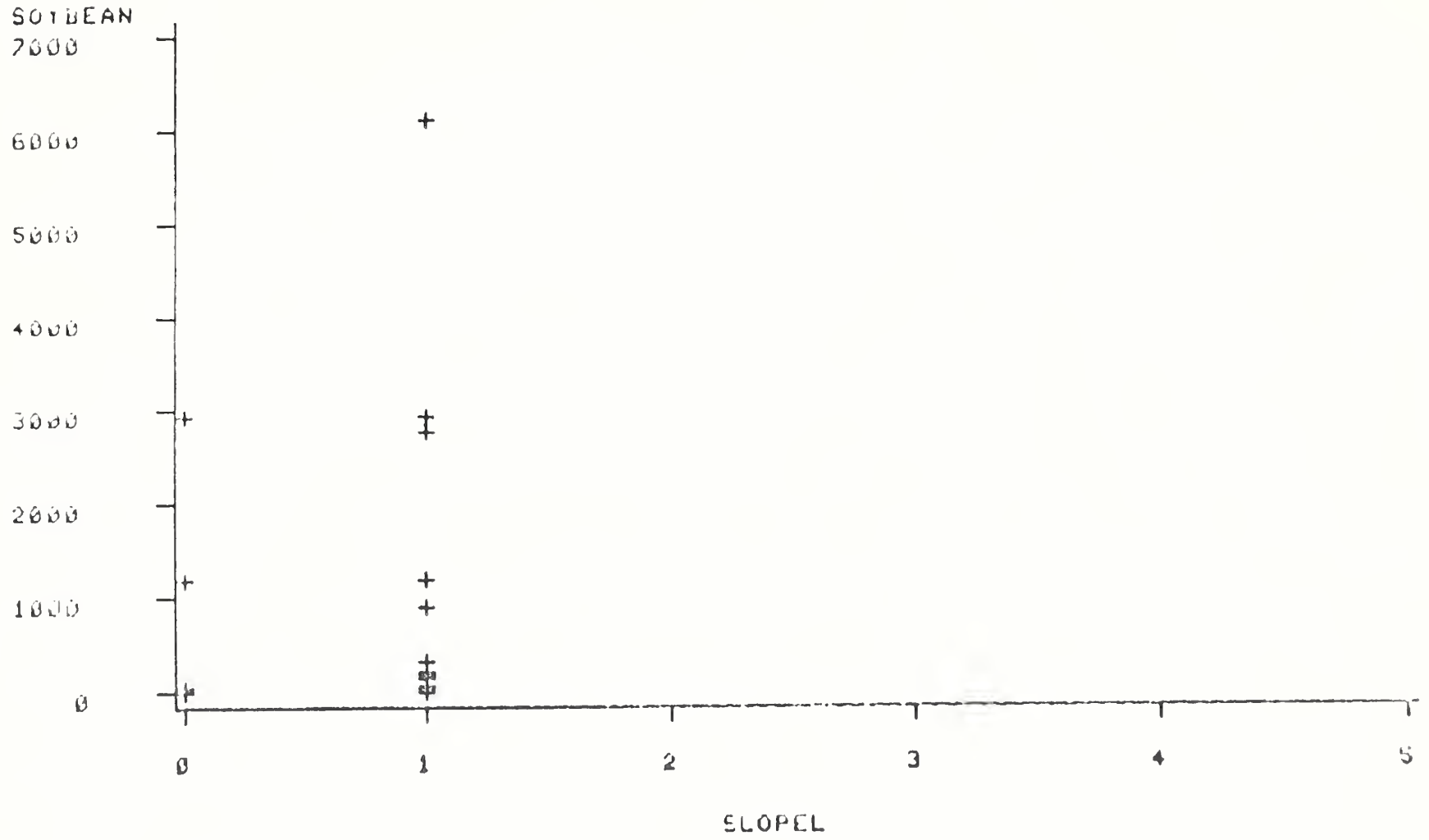
WHEAT VS. DUMMY SALT



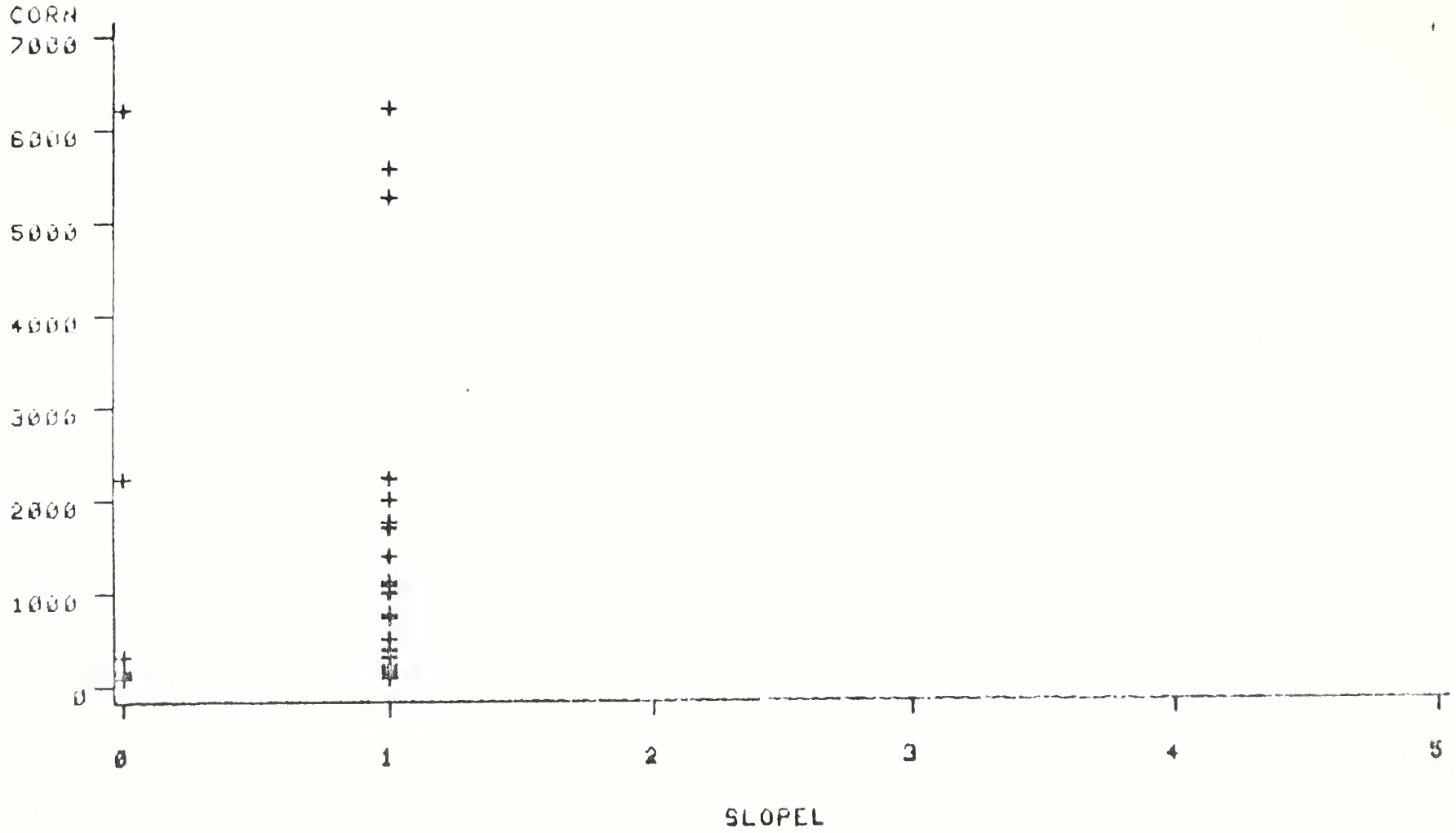
SORGHUM VS. DUMMY SLOPE



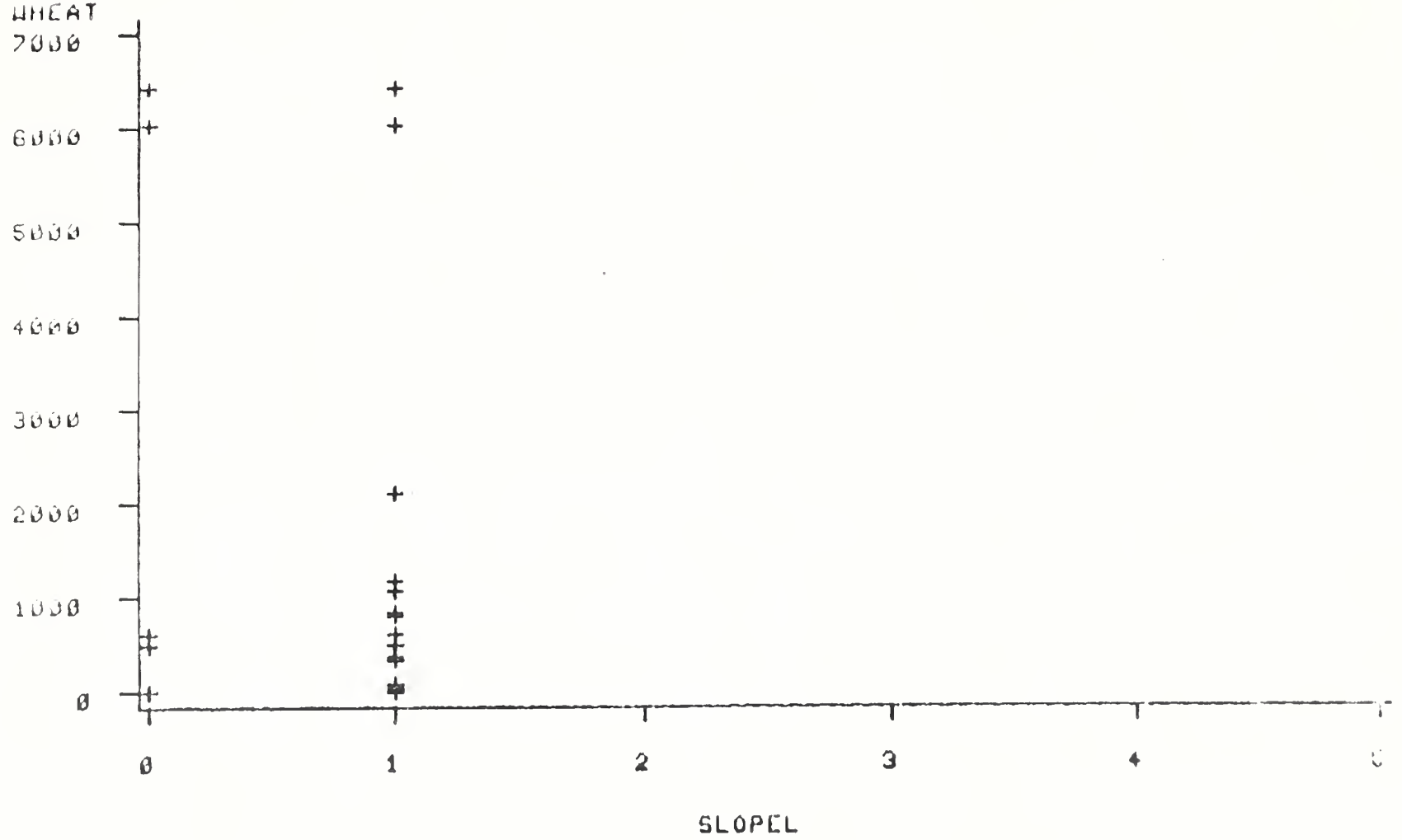
SOYBEAN VS. DUMMY SLOPE



CORN VS. DUMMY SLOPE



WHEAT VS. DUMMY SLOPE



APPENDIX 2

APPORTIONING EQUATIONS

1. County-level statistics equally apportioned to the grid cell quadrant

$$QPPSCROP = CNTYCROP * (1/CNTYOBS)$$

where: QPPSCROP is the quadrant level crop area for each of the four crops;

CNTYCROP is the crop area for each of the crops in the county;

and

CNTYOBS is the number of quadrants in the county.

2. County-level statistics apportioned using ag density

$$QPPWCROP = CNTYCROP * (W/CNTYSUMW)$$

where: QPPWCROP is the quadrant level crop area for each of the four crops;

CNTYCROP is the crop area for each of the crops in the county;

W is the ag density in each quadrant;

and

CNTYSUMW is the total ag density for all quadrants in the county.

3. State-level statistics apportioned using ag density

$$ESTCROP = STATCROP * (W/STATSUMW)$$

where: ESTCROP is the quadrant level crop area for each of the four crops;

STATCROP is the crop area for each of the crops in the state;

W is the ag density in each quadrant;

and

STATSUMW is the total ag density for all quadrants in the state.